CANADIAN THESES ON MICROFICHE

I.S.B.N.

THESES CANADIENNES SUR MICROFICHE

National Library of Canada
Collections Development Branch

Canadian Theses on
Microfiche Service

Ottawa, Canada
K1A 0N4

Bibliothèque nationale du Canada
Direction du développement des collections

Service des thèses canadiennes
sur microfiche

NOTICE

The quality of this microfiche is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us a poor photocopy.

Previously copyrighted materials (journal articles, published tests, etc.) are not filmed.

Reproduction in full or in part of this film is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30. Please read the authorization forms which accompany this thesis.

AVIS

La qualité de cette microfiche dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de mauvaise qualité.

Les documents qui font déjà l'objet d'un droit d'auteur (articles de revue, examens publiés, etc.) ne sont pas microfilmés.

La reproduction, même partielle, de ce microfilm est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30. Veuillez prendre connaissance des formules d'autorisation qui accompagnent cette thèse.

THIS DISSERTATION
HAS BEEN MICROFILMED
EXACTLY AS RECEIVED

LA THÈSE A ÉTÉ
MICROFILMÉE TELLE QUE
NOUS L'AVONS REÇUE

NL-339 (r. 82/08)

Canada

106

0-315-16015-2

National Library of Canada | Bibliothèque nationale du Canada

Canadian Theses Division | Division des thèses canadiennes

Ottawa, Canada
K1A 0N4

63924

PERMISSION TO MICROFILM — AUTORISATION DE MICROFILMER

• Please print or type — Écrire en lettres moulées ou dactylographier

Full Name of Author — Nom complet de l'auteur

WILLIAM MICHAEL CHELIAK

Date of Birth — Date de naissance	Country of Birth — Lieu de naissance
22/08/55	Canada.

Permanent Address — Résidence fixe

12 Mueller Cres.
Petawawa Ontario
K8H 3E9

Title of Thesis — Titre de la thèse

Temporal aspects of the mating system of Pinus banksiana Lamb.

University — Université

University of Alberta

Degree for which thesis was presented — Grade pour lequel cette thèse fut présentée

PhD

Year this degree conferred — Année d'obtention de ce grade	Name of Supervisor — Nom du directeur de thèse
1983	Dr. Bruce P. Dancik

Permission is hereby granted to the NATIONAL LIBRARY OF CANADA to microfilm this thesis and to lend or sell copies of the film.

The author reserves other publication rights, and neither the thesis nor extensive extracts from it may be printed or otherwise reproduced without the author's written permission.

L'autorisation est, par la présente, accordée à la BIBLIOTHÈQUE NATIONALE DU CANADA de microfilmer cette thèse et de prêter ou de vendre des exemplaires du film.

L'auteur se réserve les autres droits de publication; ni la thèse ni de longs extraits de celle-ci ne doivent être imprimés ou autrement reproduits sans l'autorisation écrite de l'auteur.

Date	Signature
20 January 1983	William M. Cheliak

NL-91 (4/77)

THE UNIVERSITY OF ALBERTA

TEMPORAL ASPECTS OF THE MATING SYSTEM OF *Pinus banksiana* Lamb.

by

WILLIAM M. CHELIAK

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES AND RESEARCH

IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE

OF DOCTOR OF PHILOSOPHY

DEPARTMENT OF FOREST SCIENCE

EDMONTON, ALBERTA

SPRING, 1983

MICRO

THE UNIVERSITY OF ALBERTA

RELEASE FORM

NAME OF AUTHOR WILLIAM M. CHELIAK

TITLE OF THESIS TEMPORAL ASPECTS OF THE MATING SYSTEM OF *Pinus banksiana* Lamb.

DEGREE FOR WHICH THESIS WAS PRESENTED DOCTOR OF PHILOSOPHY

YEAR THIS DEGREE GRANTED SPRING, 1983

Permission is hereby granted to THE UNIVERSITY OF ALBERTA LIBRARY to reproduce single copies of this thesis and to lend or sell such copies for private, scholarly or scientific research purposes only.

The author reserves other publication rights, and neither the thesis nor extensive extracts from it may be printed or otherwise reproduced without the author's written permission.

(SIGNED) William M. Cheliak

PERMANENT ADDRESS:

12 Mueller Crescent

Petawawa, Ontario, Canada

K8H 3E9

DATED 18 January 1983

THE UNIVERSITY OF ALBERTA

FACULTY OF GRADUATE STUDIES AND RESEARCH

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies and Research, for acceptance, a thesis entitled Temporal aspects of the mating system of *Pinus banksiana* Lamb. submitted by WILLIAM M. CHELIAK in partial fulfilment of the requirements for the degree of DOCTOR OF PHILOSOPHY in FOREST SCIENCE.

Bruce P. Dancik
.....
Supervisor

Kenneth D. Higginbotham
.....

Kenneth Morgan
.....

Curtis Strobeck
.....

Francis C. Yeh
.....

Anthony Brown
.....
External Examiner

Date 18 June 1982

This thesis is dedicated to Leanne, Jeff and Paul

Abstract

Segregation ratios and parameters of the mating system of jack pine (*Pinus banksiana* Lamb.) were studied in a stand near Bruderheim, Alberta. Stability of these parameters was assessed for three crown strata and four independent fertilization events within the canopy of this stand using open-pollinated seeds from serotinous cones harvested in 1980 from thirty trees.

Analyses of observed segregation ratios of four of five polymorphic isozyme loci showed good correspondence to the 1:1 ratios expected under simple Mendelian control; Allozymes of glucose-6-phosphate dehydrogenase did not segregate in the expected 1:1 ratio. However, there were significant segregation distortions at some sampling positions on individual trees. Heterogeneity among trees, strata and years could be the result of pollen pool heterogeneity, segregation distortion and/or primitive lethal and semi-lethal gene combinations resulting in early embryo abortion. These types of segregation distortion are important in estimation of population allele frequencies from bulked samples of a small number of individuals as well as in inference of heterozygosity of individual plants.

The mixed mating system model with a finite number of effective males accurately described the mating system of jack pine. The average outcrossing rate was estimated to be 88±.047%. However, there were significant differences in the mating system among three of the four years; there was an

apparent linear decline of the effective selfing rate with age of the seed. Real changes in the mating system as well as higher mortality of selfed-embryos over time could account for these observations. No differences could be demonstrated among the three crown strata of this stand. There was significant inter-locus heterogeneity in the filial genotypic distributions and in the estimated outcrossing rates, reflecting the complex nature of forces that can affect single locus estimates. Accurate estimates of the mating system, therefore, should be based upon a large number of loci. There was evidence of some additional inbreeding due to family structures in the stand; however, this was a minor component of the total inbreeding. These results have implications in applied breeding programs for estimation of genetic parameters and inference of genetic differences from open-pollinated progeny arrays.

Acknowledgements

I would like to express my sincere appreciation and gratitude to all of the members of my graduate committee for their assistance, persistence and encouragement during this study. I would also like to express my thanks to Ms. H. Klemm and Mr. J. Lavoie for excellent assistance with the laboratory portion of this study. A special thanks goes to Dr. Narinder Dhir of the Alberta Forest Service for many fruitful and interesting discussions as well as in assistance with logistical problems of the field work. Thanks to Mr. John King for assistance with the field work. Many thanks to Ms. M. Holmes for assistance with many of the last minute scrambling to put it all togeather.

Funds provided by the Natural Sciences and Engineering Research Council of Canada and the Canadian Forestry Service (Petawawa National Forestry Institute) are gratefully acknowledged.

Table of Contents

Chapter Page

1. INTRODUCTION AND OBJECTIVES 1

1.1 Literature Cited 6

2. INHERITANCE, LINKAGE AND SEGREGATION ANALYSIS 10

2.1 Introduction 10

2.2 Materials and Methods 12

2.3 Results 18

2.4 Acid phosphatase (Aph) E.C. 3.1.3.2 19

2.5 Malate dehydrogenase (Mdh) E.C. 1.1.1.37 22

2.5.1 Mdh-3 25

2.5.2 Mdh-4 28

2.6 6-Phosphogluconic dehydrogenase (6Pgd) E.C. 1.1.1.44 31

2.6.1 6Pgd-2, allele 2 32

2.6.2 6Pgd-2, allele 4 36

2.7 Glucose-6-phosphate dehydrogenase (G6p) E.C. 1.1.1.49 36

2.8 Additional Variable Loci 42

2.8.1 Aspartate aminotransferase (Aat-1) E.C. 2.6.1.1 42

2.8.2 Alcohol dehydrogenase (Adh) E.C. 1.1.1.1 ..45

2.8.3 Glutamate dehydrogenase (Gdh) E.C. 1.4.1.2 46

2.8.4 Malic enzyme (Me) E.C. 1.1.1.40 46

2.9 Linkage Analaysis 47

2.10 Discussion 48

2.11 Literature Cited 53

3. ESTIMATION OF MATING SYSTEM AND EFFECTIVE POPULATION SIZES 60

3.1 Introduction 60

3.2 Materials and Methods 61

3.3 Results 64

3.4 Filial population 66

3.4.1 Loci in Hardy-Weinberg equilibrium 66

3.4.2 Loci not in Hardy-Weinberg equilibrium 68

3.4.3 Summary of levels of variation 68

3.5 Estimation of the mating system 70

3.5.1 Comparison of methods 70

3.5.2 Mating system of jack pine 74

3.5.3 Log-linear analyses 74

3.5.4 Multilocus analysis 79

3.6 Pollen pool 80

3.6.1 Comparison of the mature and outcrossed allele frequencies 80

3.6.2 Homogeneity of the pollen pool 82

3.7 Variance effective population numbers 83

3.8 Discussion 86

3.9 Literature Cited 95

4. PRACTICAL IMPLICATIONS 100

4.1 Introduction 100

4.1.1 Inference 100

4.1.2 Estimation 101

4.2 Seed Orchards 106

4.2.1 Current nursery,greenhouse and plantation practices 107

4.3 Literature Cited 108

4.3 Literature Cited 108
5. APPENDIX 1. 110
6. APPENDIX 2. 113
6.1 The expectation step (E) 114
6.1.1 The maximization step (M) 117
6.2 Literature cited 119

List of Tables

Table		Page
2.1	Inheritance/linkage studies in some forest trees.	11
2.2	Gel and electrode buffers and conditions used for the study.	16
2.3	Staining recipies and running conditions used in the inheritance and linkage study for jack pine.	17
2.4	Observed segregation of allele 1 and 2 of enzyme Aph-2 by female, crown strata and year of collection.	21
2.5	Summary of the segregation analysis for Aph-2 by female, crown strata and year of collection.	23
2.6	Observed segregation of allele 1 and 2 of enzyme Mdh-3, by female, crown strata and year of collection.	26
2.7	Summary of the segregation analysis of Mdh-3, by female, crown strata and year of collection.	27
2.8	Observed segregation of allele 1 and 2 of enzyme Mdh-4 by female, crown strata and year of collection.	29
2.9	Summary of the segregation analysis of Mdh-4 by female, crown strata and year of collection.	30
2.10	Observed segregation of allele 1 and 2 of enzyme 6Pgd-2 by female, crown strata and year of collection.	33

2.11 Summary of the segregation analysis of 6Pgd-2 allele 2 by female, crown strata and year of collection. 35

2.12 Observed segregation of allele 1 and 4 of enzyme 6Pgd-2 by female, crown strata and year of collection. 37

2.13 Summary of the segregation analysis for 6Pgd-2, allele 4 by female, position and year of collection. 38

2.14 Observed segregation of allele 1 and 2 of enzyme G6p by female, crown strata and year of collection. 41

2.15 Summary of the segregation analysis for G6p by female crown strata and year of collection. 43

3.1 Allele frequencies, heterozygosities and test results for Hardy-Weinberg equilibrium distributions in the mature and filial populations. 65

3.2 Summary of inbreeding coefficients (F) calculated from equation 2 for year and crown stratum subdivisions in the filial populations. 67

3.3 Estimates of the mating system of jack pine using the methods of Brown *et al.* (1975) (=B) and Cheliak *et al.* (1982) (=EM). 71

3.4 ANOVA of logit of expected numbers of outcrossed versus selfed progeny by log-linear models. 75

3.5 G^2 summary of homogeneity of outcrossing rates for single loci and pooled summaries. 77

3.6 t-test for the homogeneity of common allele frequencies observed in the mature population (female component) and the observed pollen pool (male component). 81

3.7 Variance effective population numbers and variance effective male numbers calculated from the mature population and observed pollen pool. 84

4.1 Expected coefficients of additive genetic variance under inbreeding and various family structures for the estimated mating system of a stand of jack pine. 104

A2.1 Observed phenotypes and underlying genotypes resulting from various mating events. 115

List of Figures

Figure		Page
2.1	Geographic location of the population sampled and the distribution of the trees sampled within the stand. The numbers refer to the sample number assigned to each tree. The stand is located near Bruderhiem, Alberta.	46
2.2	Diagram of acid phosphatase gels indicating haploid megagametophytes in lanes 1 to 4 and diploid embryo phenotypes in lanes 5 to 8. Lane: 5 = 22, 6 = 12, 7,8 = 11	20
2.3	Diagram of malate dehydrogenase gels indicating haploid megagametophytes in lanes 1 to 4 and diploid embryo phenotypes in lanes 5 to 8. Mdh-3 Lane: 5 = 12, 6 = 22, 7,8 = 11. Mdh-4 Lane: 5 = 22, 6 = 12, 7 = 11, 8 = 12.	24
2.4	Diagram of 6-phosphogluconic dehydrogenase gel indicating haploid megagametopytes in lanes 1 to 4 and diploid embryo phenotypes in lanes 5 to 9. Lane: 5 =12, 6 = 22, 7 = 11, 8 = 14, 9 = 13.	32
2.5	Diagram of glucose-6-phosphate dehydrogenase gel indicating haploid megagametophytes in lanes 1 to 4 and diploid embryo phenotypes in lanes 5 to 8. Lane: 5 = 11, 6 = 22, 7 = 12, 8 = 13.	40
2.6	Diagram of aspartate aminotransferase gel indicating haploid megagametophytes in lanes 1 to 4 and diploid embryo phenotypes in lanes 5 to 8. Lane: 5 = 11, 6 = 12, 7 = 13, 8 = 14.	44

2.7 Diagram of alcohol dehydrogenase and glutamate dehydrogenase gel indicating haploid megagametophytes in lanes 1 to 4 and diploid embryo phenotypes in lanes 5 to 8. Adh Lane: 5 = 11, 6 = 12, 7 = 11, 8 = 11. Gdh Lane: 5 = 11, 6 = 11, 7 = 12, 8 = 11. 46

2.8 Diagram of malic enzyme gel indicating haploid megagametophytes in lanes 1 to 4 and diploid embryo phenotypes in lanes 5 to 8. Lane: 5 = 11, 6 = 12, 7 = 13, 8 = 12. 47

4.1 Expected coefficent of additive genetic variance (K^{-1}) with various levels of inbreeding (F). 103

1. INTRODUCTION AND OBJECTIVES

The essential connection between any two generations of a sexually reproducing species is the pattern by which gametes are transmitted from the male to the female. This pattern of transmission is usually termed the mating system of a species. Thus, the mating system defines a functional relationship which allows us to predict how the gametes from the mature population combine to form the viable zygotic population in a sample.

Plants possess several important characteristics, which make them attractive organisms for the study of mating systems. Among these are the great variety of mating systems (Fryxell 1957), ranging from highly selfed to highly outcrossed species, production of large numbers of progeny per maternal plant, and the relative ease with which these progeny can be collected, stored and grown for subsequent analysis.

Traditional approaches to the analysis of mating systems have been largely qualitative and involved observations of floral morphology, crossing experiments and, where appropriate, pollen movement and pollinator behavior. These qualitative studies resulted in classification systems which viewed species, for example, as predominantly selfed or outcrossed or as self-compatible or incompatible. The experiments by D.F. Jones (1916, 1928) were notable exceptions to this approach. These were some of the earliest investigations on the dynamic nature of plant mating

systems, where it was demonstrated that species were not wholly selfed or outcrossed, but instead exhibited a range between these two extremes, dependent upon many external and internal mechanisms. It was not until the early 1950's that population geneticists began to more fully appreciate the dynamic nature of plant mating systems, due largely to Fyfe and Bailey (1951), who provided maximum-likelihood methods to estimate pollen allele frequency (p) and outcrossing rate (t) simultaneously from population data. Therefore, it was no longer necessary to design experiments where two alternate homozygous genotypes were the only individuals participating in mating events. Given suitable genetic markers, it was possible with this procedure to make estimates of the mating system of species in natural populations or in field plantations.

Subsequent to this initial breakthrough, a large number of estimators were developed to estimate mating system parameters in natural populations. These estimators and their relative merits and shortcomings have been summarized by Vasek (1973) and Jain (1979). However, the dominant/recessive expression of most characters studied necessitated extensive progeny testing to identify heterozygotes.

The application of protein electrophoresis resulted in a significant advance in the estimation of mating system parameters. Since essentially all allozymes are expressed co-dominantly, all phenotypes of a locus can be observed

without progeny testing. An additional advantage of this technique is the large number of markers that are available for multiple estimates of mating system parameters. In response to application of electrophoretic techniques to the study of mating systems, there also has been research into suitable estimation procedures. Numerous single-locus maximum-likelihood estimators (Brown and Allard 1970, Brown *et al.*, 1975 Clegg *et al.* 1978, Cheliak *et al.* 1982) as well as multiple-locus estimators (Green *et al.* 1980, Ritland and Jain 1981, Shaw *et al.* 1981) now exist to estimate mating system parameters in natural or experimental populations.

While most of our knowledge on mating systems has been obtained from short-generation plants, some estimates of the mating system have been made in typically long-generation forest tree species, using quantitative (Squillace 1974) and electrophoretic markers (Adams and Joly 1980, Brown *et al.* 1975, El-Kassaby *et al.* 1982, Hopper and Moran 1981, Moran and Brown 1980, Moran *et al.* 1981, Phillips and Brown 1977, Yeh *et al.* 1982). In gymnosperms, where double fertilization does not occur, electrophoretic techniques are especially powerful tools to estimate mating system parameters. The lack of double-fertilization results in a haploid endosperm (megagametophyte). Thus, in addition to direct probabilistic inference of maternal genotypes by analysis of haploid tissues from one plant (Morris and Speith 1978) inheritance and linkage studies can be made simultaneously without the necessity of performing crosses. Furthermore, since all of

the cells within the seed have been derived mitotically from a single haploid megaspore, the maternal haploid contribution to the corresponding embryo can be unambiguously deduced directly from the megagametophyte. When the haploid maternal genotype and corresponding diploid embryo genotype are known, the paternal contribution can then be unambiguously ascertained. Thus, progeny arrays from seeds of conifers contain considerable information which can be used to study mating systems.

Jack pine (*Pinus banksiana* Lamb.) is a monoecious conifer typical of boreal forests in the northern United States and Canada. Its natural range extends from the Atlantic coast of Maine and Nova Scotia through the Great Lakes Region, across the northern prairies to the valley of the Mackenzie River (Fowells 1965). Due to its low shade tolerance and serotinous cones, jack pine typically regenerates naturally after fire has destroyed the parental stems and has provided sufficient heat to open the serotinous cones. This type of regenerative process usually results in relatively even-aged, dense stands.

Geographic and genetic variation has been noted in climatic adaptation and virtually all characteristics related to growth and form of jack pine (Rudolph and Yeatman 1982). Commercially, the species is important for a number of products within its natural range, and tree improvement programs have been established in several regions. Selection of breeding stock from wild stands and domestication

programs for breeding purposes, however, are in a relatively early stage of development. Estimates of the mating system in natural populations could provide insights into past evolutionary forces (such as evolution of inbreeding) and could be incorporated at an operational level in applied breeding programs. Thus, the study of mating systems can make significant contributions to the understanding and effective manipulation of genetic variation in a species.

The purpose of this study was to estimate the mating system of jack pine. The specific objectives were:

1. To determine the inheritance and linkage patterns of a number of polymorphic isozyme markers in jack pine by the technique of electrophoresis using haploid megagametophtyes and viable embryos. A first and necessary step to utilize electrophoretic markers effectively is to demonstrate that they are inherited in a simple Mendelian fashion.
2. To document the mating system and related parameters of jack pine. This part of the study centered around estimating parameters of the mating system of jack pine as well as the effective population size of the parental generation from which the seeds were derived and the homogeneity of the pollen pool. Some practical implications of the results to breeding programs are discussed.

The serotinous cones of jack pine provide a convienient way to study temporal variation in segregation ratios and

parameters of the mating system. Previous studies on pollen flow dynamics of wind pollinated species (Koski 1970) and distribution of male and female flowers in conifers (Sarvas 1962) suggest that there may be crown stratum differences in the genetic composition of the pollen pool. Therefore, this study assesses differences in the mating system and segregation ratios for various levels of the reproductive crown in addition to temporal differences.

The results of this work have been organized into two central chapters. The first deals with inheritance and linkage of the polymorphic loci as well as the stability of the observed segregation ratios. The second chapter presents estimates of the mating system, pollen pool homogeneity, and variance effective population sizes for various temporal and spatial dimensions. The final chapter discusses the practical implications and implementation of the mating system results in an applied breeding program.

1.1 Literature Cited

Adams, W.T., and R.J. Joly. 1980. Allozyme studies in loblolly pine seed orchards: Clonal variation and frequency of progeny due to self-fertilization. Silvae Genetica 29:1-4.

Brown, A.H.D., and R.W. Allard. 1970. Estimation of the mating system in open-pollinated maize populations using isozyme polymorphisms. Genetics 66:133-145.

Brown, A.H.D., A.C. Matheson and K.G. Eldrige. 1975. Estimation of the mating system of *Eucalyptus obliqua* L'Herit. using allozyme polymorphisms. Aust. J. Bot. 23:931-949.

Cheliak, W.M., K. Morgan, C. Strobeck, F.C.H. Yeh and B.P. Dancik. 1982. Estimation of mating system parameters in plant populations using the EM algorithm. (Submitted Theor. Appl. Genet.)

Clegg, M.T., A.L. Kahler and R.W. Allard. 1978. Estimation of life cycle components of selection in an experimental plant population. Genetics 89:765-792.

El-Kassaby, Y.A., F.C. Yeh and O. Sziklai. 1982. Estimation of the outcrossing rate using allozyme polymorphisms in a natural stand of coastal Douglas-fir *Pseudotsuga menziessii* var. *menziesii* Mirb. Franco.) (In press Silvae Genetica)

Fowells, H.A. 1965. Silvics of forest trees of the United States. U.S.D.A For. Serv. Agric. Handbook 271.

Fyfe, J.L., and N.T.J. Bailey. 1951. Plant breeding studies in *Leguminous* forage crops. I. Natural cross-breeding in winter beans. J. Agri. Sci. 41:371-378.

Fryxell, P.A. 1957. Mode of reproduction in higher plants. Bot. Rev. 23:135-233.

Green, A.G., A.H.D. Brown and R.N. Oran. 1980. Determination of outcrossing rate in a breeding population of *Lupinus albus* L. (white lupin). Z. Pflanzenzuchtg. 84:181-191.

Hopper, S.D., and G.F. Moran. 1981. Bird pollination and the

mating system of *Eucalyptus stoatei*. Aust. J. Bot. 29:625-638.

Jain, S.K. 1979. Estimation of outcrossing rates: Some alternative procedures. Crop Sci. 19:23-26.

Jones, D.F. 1916. Natual cross-pollination in the tomato. Science 43:509-510.

Jones, D.F. 1928. Selective fertilization. Univ. of Chicago Press, Chicago.

Koski, V. 1970. A study of pollen dispersal as a mechanism of gene flow in conifers. Comm. Inst. For. Fenn. 70(4):1-78.

Moran, G.F., and A.H.D. Brown. 1980. Temporal heterogeneity of outcrossing rates in alpine ash (*Eucalyptus delegatensis* R.T. Bak.). Theor. Appl. Genet. 57:101-105.

Moran, G.F., J.C. Bell and A.C. Matheson. 1980. The genetic structure and levels of inbreeding in a *Pinus radiata* D. Don seed orchard. Silvae Genetica 29:190-193.

Morris, R.W., and P.T. Speith. 1978. Sampling strategies for using female gametophytes to estimate heterozygosity in conifers. Theor. Appl. Genet. 51:217-222.

Phillips, M.A., and A.H.D. Brown. 1977. Mating system and hybridity in *Eucalytpus pauciflora*. Aust. J. Biol. Sci. 30:337-344.

Ritland, K., and S.K. Jain. 1981. A model for the estimation of outcrossing rate and gene frequencies using *n* independent loci. Heredity 47:35-52.

Rudolph, T.D., and C.W. Yeatman. 1982. Genetics of jack

pine. U.S.D.A. For. Serv. Publ. (in press).

Sarvas, R. 1962. Investigations on the flowering and seed crop of *Pinus silvestris*. Comm. Inst. For. Fenn. 53(4):1-198.

Shaw, D.V., A.L. Kahler and R.W. Allard. 1981. A multilocus estimator of mating system parameters in plant populations. Proc. Nat. Acad. Sci. (USA) 78:1298-1302.

Squillace, A.E. 1974. Average genetic correlations among offspring from open-pollinated forest trees. Silvae Genetica 23:149-156.

Vasek, F.C. 1973. Outcrossing in natural populations: A comparison of outcrossing estimation methods. p 369-385. *In* Evolution and Environment (E.T Drake, ed.) Yale Univ. Press, New Haven.

Yeh, F.C.H., A. Brune, W.M. Cheliak and D.C. Chipman. 1982. Mating system of *Eucalyptus citriodora* in a seed stand. (Submitted Can. J. For. Res.).

2. INHERITANCE, LINKAGE AND SEGREGATION ANALYSIS

2.1 Introduction

The utility of electrophoresis for studying gene structure, genetic differentiation and mating systems of populations has been shown for a large number of diverse organisms (Allard 1975, Brown 1979, Nevo 1978, Lewontin 1974, Rick *et al.* 1977, Shaw *et al.* 1981). However, before electrophoresis can be used to its fullest advantage in these types of studies, inheritance of the electrophoretically detectable polymorphisms as well as their linkage relationships must be known.

Inheritance and linkage have been studied in a number of coniferous species (Table 2.1). The nature of the life cycle in conifers permits inheritance and linkage studies to be performed simultaneously. Unlike angiosperms, the endosperm (female gametophyte or megagametophyte) of gymnosperms is haploid in the mature seed. This condition allows the genotype of an individual to be determined probabilistically by analysis of haploid megagametophytes without the necessity of performing crosses. Furthermore, since one haploid megaspore mitotically gives rise to all other cells within the ovule, barring mutation, the embryo will inherit a haploid genotype identical to the megagametophye. In this chapter, the inheritance of five polymorphic enzyme systems of megagametophytes of jack pine

Table 2.1. Inheritance/linkage studies in some forest trees.

Species Name	# of Enzymes	# of Loci	Type	Reference
loblolly pine	10	17,12	I,L	Adams and Joly 1980a b
loblolly pine	10	12	L	Adams and Joly 1980b
Norway spruce	1	1	I	Bartels 1971
Norway spruce	2,3	2,6	I,L	Bergmann 1973 1974
knobcone pine	2	5	I	Conkle 1971
Douglas-fir	5,7	8,9	I,L	El-Kassaby *et al* 1982a b
white pine	10	17	I,L	Eckert *et al.* 1981
pitch pine	8	14	I	Guries and Ledig 1978
pitch pine	7	11	L	Guries *et al.* 1978
Norway spruce	1	1,2	I	Lundkvist 1974 1975 1977
Norway spruce	3	4	I	Lundkvist, Rudin 1977
Norway spruce	5	11	I	Lundkvist 1979
ponderosa pine	12	23	I,L	O'Malley *et al.* 1979
Scots pine	1	2	I,L	Rudin 1975 1977
Scots pine	6	12	L	Rudin and Ekberg 1978
jack pine	7	7	I	Tobolski 1979
black spruce	12	18	I,L	Yeh *et al.* 1982

I = inheritance study; L = linkage study

(*Pinus banksiana* Lamb.) is described. In addition, open-pollinated progeny arrays were assayed to determine if the polymorphic enzyme loci surveyed were inherited in a simple Mendelian fashion.

2.2 Materials and Methods

Thirty sexually mature jack pine trees were sampled from a stand near Bruderheim, Alberta (Figure 2.1). To reduce the possibility of sampling closely related individuals (i.e., full or half siblings), a minimum of 30 m spacing between sample trees was maintained. To assess temporal and spatial variation, the active female reproductive crown of each tree was arbitrarily stratified into three strata: low (L), middle (M) and high (H). Serotinous cones representing pollinations of 1975, 1976, 1977 and 1978 were harvested from each stratum of each tree from a randomized cardinal direction in the late fall of 1980. Thus, there were 12 sampling positions per tree.

Open-pollinated seeds were extracted from the serotinous cones by the following method (A.K. Hellum pers. comm.): Cones were placed in an oven at 180° C for two minutes to break the resin bonds of the cone scales. Cones were then placed in an oven at 60° C for 8-10 hours to reflex more fully the ovuliferous scales. The seeds were then removed from the cones by lightly tapping the cones on a hard surface.

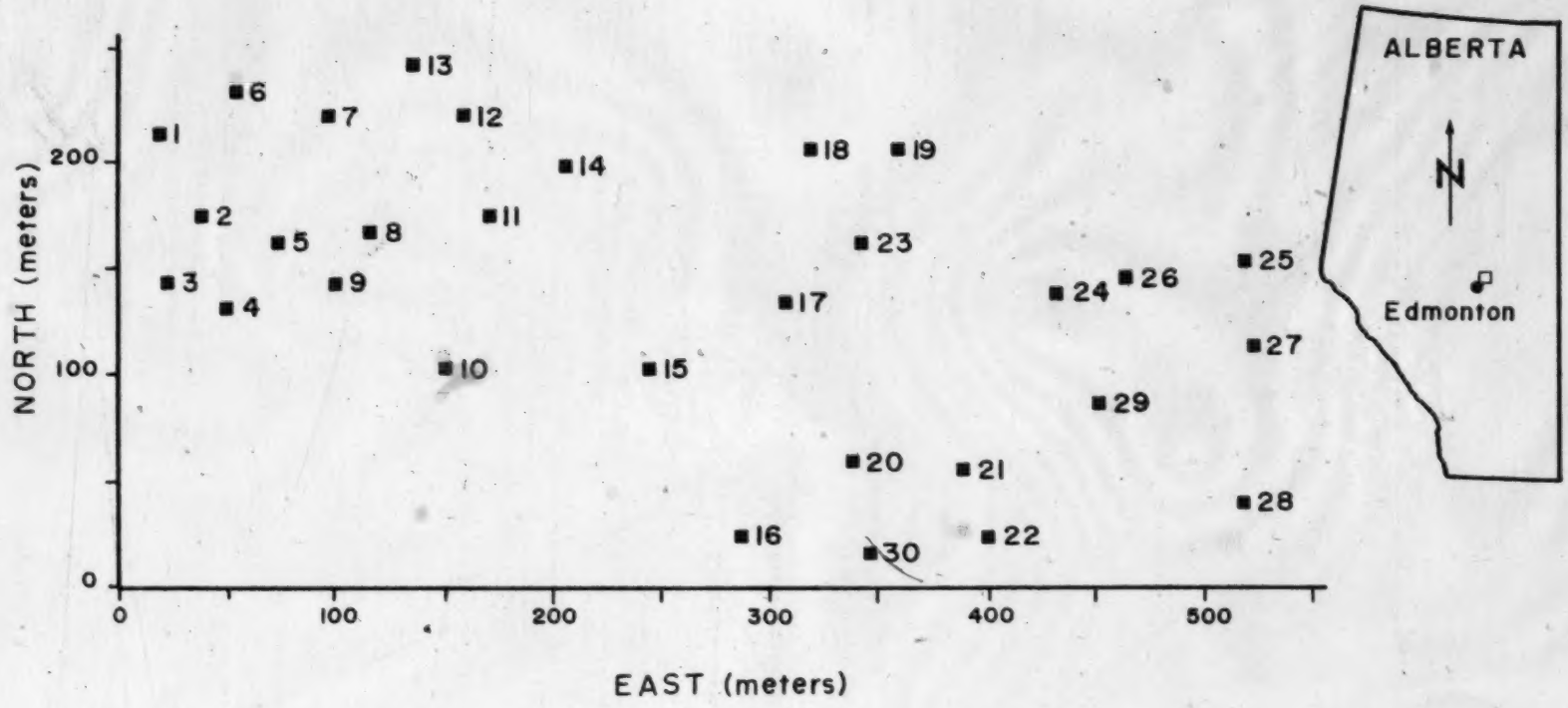

Figure 2.1. Geographical location of the population sampled and the distribution of the trees sampled within the stand. The numbers refer to the sample number assigned to each tree. The stand (□) is located near Bruderheim, Alberta.

Since detection of a heterozygote was of primary concern, required sample sizes were calculated by:

$$P_h = 1 - (.5)^{n-1} \qquad -1-$$

where n is the number of gametes and P_h is the probability of detecting a heterozygote, given that it is a heterozygote. This sampling strategy assumes equal distribution and survival of alternate gametes in a heterozygous individual. In addition, the open-pollinated progeny (i.e., embryos) allow testing of the hypothesized maternal genotype under the assumptions of a single co-dominant, fully penetrant gene locus segregating in a simple Mendelian fashion.

For maternal genotyping, seeds were allowed to imbibe distilled water for 24 hours and either were assayed immediately or refrigerated for up to one week before assay. Embryos for the progeny arrays were germinated for about 72 hours before being assayed. There was, however, a great deal of inter-family variation in 72-hour germination of the seeds. Therefore, it was decided on the basis of a preliminary ontogenetic study to allow the radicle to emerge approximately 2-3 mm from the seed before embryo assay.

Seed coats and embryos were removed for maternal genotyping. Individual megagametophytes were crushed in a scintillation cup (0.5 ml) with one drop of extraction buffer (Yeh and O'Malley 1980). The crude homogenate was

absorbed on three, 10x1 mm filter paper (Watmann No. 3) wicks. Up to 40 megagametophytes were assayed per tree for the preliminary genotyping for the first experiment. Individual embryos, with 2-3 mm of radicle, were prepared in the same manner. The female gametophyte and the corresponding embryo were placed at adjacent positions in the gels. Generally, 10 embryos were assayed per sampling position (12 positions per tree).

The wicks with the crude homogenate were inserted into a slice (the origin) in the gel, 2 cm from the cathodal end. 12.5% (wt/vol) starch gels were prepared from Electrostarch (Electrostarch Co., Madison WI, lot 307). The components of the gel and electrode buffers are shown in Table 2.2.

Electrophoresis was performed under refrigeration at approximately 2° C. The proteins were introduced into the gel at 1/2 running voltage for 10 minutes. Wicks were then removed, crushed ice packs were applied to the upper surface of the gel and total running voltage was applied (Table 2.2). Electrophoresis was conducted until the borate front (buffer system LC) or the red food coloring tracker dye (buffer systems MC and TC) had migrated a minimum of 5 cm beyond the origin. Following electrophoresis, gels were cut into 1 mm thick horizontal slices, stained, and incubated at 37° C (Table 2.3).

A maximum likelihood G^2 (asymptotic X^2) was used to test for 1:1 segregation of alleles in the megagametophytes. An heterogeneity G^2 was calculated to test the homogeneity

Table 2.2 Gel and electrode buffers and conditions used for the study.

Morpholine-Citrate buffer system (MC)

Electrode Buffer: 0.04 M citric acid (anhydrous), adjust pH to 6.1 with N-(3-aminopropyl)-morpholine.

Gel Buffer: 1:20 dilution of electrode buffer.

Power: 180 V (ca. 40 mA).

Reference: Clayton and Tretiak (1972).

Lithium Citrate buffer system (LC)

Electrode Buffer: 0.06 M lithium hydroxide, 0.3 M boric acid pH 8.1.

Gel Buffer: 0.03 M Tris, 0.005 M citric acid (anhydrous) and 1% of electrode buffer, pH 8.5.

Power: 250 V (ca. 60 mA)

Reference: Ridgeway *et. al.*, (1970).

Tris-Citrate buffer system (TC)

Electrode buffer: 0.13 M Tris and 0.043 M citric acid (anhydrous), pH 7.0.

Gel Buffer: 1:15 dilution of electrode buffer.

Power: 180 V (ca. 40 mA).

Reference: Siciliano and Shaw (1976).

Table 2.3. Staining recipies and running conditions used in the inheritance and linkage study for jack pine.

Aspartate aminotransferase (Aat) :running buffer LC; stain: 50 ml AAT substrate[2], 5 mg Pyroxidal-5' phosphate, 100 mg Fast Blue BB salt

Alcohol/Glutamate dehydrogenase (Adh/Gdh): running buffer LC; stain: 50 ml 0.2 M Tris-HCl pH 8.0, 15 mg NAD, 10 mg NBT 10, 10 mg MTT, 10 ml EtOh, 400 mg L-Glutamic acid

Acid phosphatase (Aph): running buffer MC; stain: 50 ml 0.2 M Sodium Acetate pH 5.0, 150 mg Na-α-napthyl acid phosphate, 5 ml $MgCl_2$ 1% W/V, 75 mg Fast Garnet GBC salt

Malate dehydrogenase (MDH): running buffer MC; stain: 50 ml 0.2M Tris-HCl pH 8.0, 10 mg NAD, 10 mg NBT, 5 mg PMS, 25 ml 0.5 M DL-Malic acid

Malic enzyme (ME):running buffer MC; stain: 25 ml MC E.B., 10 mg NADP, 1 ml $MgCl_2$ 1% W/V, 10 mg MTT, 5 mg PMS, 25 ml 0.5 M DL-Malic acid

Glucose-6-phosphate dehydrogenase (G6P): running buffer MC; 50 ml 0.2 M Tris-HCl pH 8.0, 10 mg NADP, 1 ml $MgCl_2$ 1% W/V, 10 ml MTT, 5 mg PMS, 100 mg Glucose-6-phosphate

Isocitric dehydrogenase (IDH): running buffer TC; stain: 50 ml 0.2 M Tris-HCl pH 8.0, 10 mg NADP, 1 ml $MgCl_2$ 1% W/V, 10 mg MTT, 5 mg PMS, 100 mg DL-Isocitric acid

6-Phosphogluconic dehydrogenase (6PG): running buffer; TC stain: 10 ml 0.2 M Tris-HCl pH 8.0, 10 mg NADP, 1 ml $MgCl_2$ 1% W/V, 15 mg PMS, 10 mg Phosphogluconic acid,

---

[1]**Buffer systems:** MC = Morpholine Citrate; LC = Lithium Citrate; TC = Tris Citrate. See Table 2 for components of each of the buffer systems.

[2]**AAT substrate:** 5.3 g L-aspartic acid, 700 mg a-Ketoglutaric acid, dissolve in 0.2 M Tris-HCl and adjust pH to 8.0 with 1 M NaOH

---

of the observed segregation ratios among females and among sampling positions within a female. A test for linkage was made for multiply heterozygous females. The test criterion was a G^2 analysis conditional on the null hypothesis of a 1:1:1:1 distribution of gametes at two heterozygous loci within a female.

A deviation from the expected ratios is judged statistically significant if it exceeds that of chance expectation at α=.05. No indication was made for levels lower than this level of significance.

2.3 Results

Five loci (Aph-2, G6p, Mdh-3, Mdh-4 and 6Pgd-2) were inferred to be segregating in this population sample from analysis of female gametophytes. Variation at additional enzyme loci was detected in aconitase (E.C. 4.2.1.3), adenelate kinase (E.C. 2.7.4.3), phosphoglucose isomerase (E.C. 5.3.1.9), 6-phosphogluconic dehydrogenase-1 (E.C. 1.1.1.44), leucine aminopeptidase (E.C. 3.4.11.1), aspartate aminotransferase-2 (E.C. 2.6.1.1), and diaphorase (NADH dependent) (E.C. 1.6.4.3) using only the haploid megagametophytic tissue. However, problems in consistently resolving these systems or in interpretation using the diploid embryos precluded their inclusion in this study. Since the major objective of this study was to assess spatial and temporal variation in the mating system of jack

pine, comparisons of segregation were made by year and crown stratum. For each variable locus, a discussion of the allozyme variants detected and detailed analysis of segregation by female, year and crown stratum follow.

2.4 Acid phosphatase (Aph) E.C. 3.1.3.2

Two zones of activity were observed on gels scored for acid phosphatase (Figure 2.2). Both zones segregated in the megagametophytes of trees heterozygous for single-banded allozyme variants. However, problems in consistently resolving the more anodal zone under the electrophoretic conditions of this study precluded its inclusion. At zone 2 (called Aph-2), embryos were either single-banded at the same position as the megagametophyte, or triple-banded with one or two bands coincident with one or both bands in the megagametophytes. In addition, an intermediate band of approximately double the intensity was observed in these embryos. This strongly suggests that Aph enzymes are functionally dimeric, as has been reported in *Picea abies* (L. (Karst.)) by Lundkvist (1977).

In the single tree heterozygous for Aph-2, overall observed segregation did not fit the expected 1:1 distribution of gametes (G_1^2=8.78, where the subscript represents the degrees of freedom; Table 2.4). There was no detectable heterogeneity among the pooled distributions for crown strata (G_1^2=0.74) or years (G_1^2 = 0.25). Only one

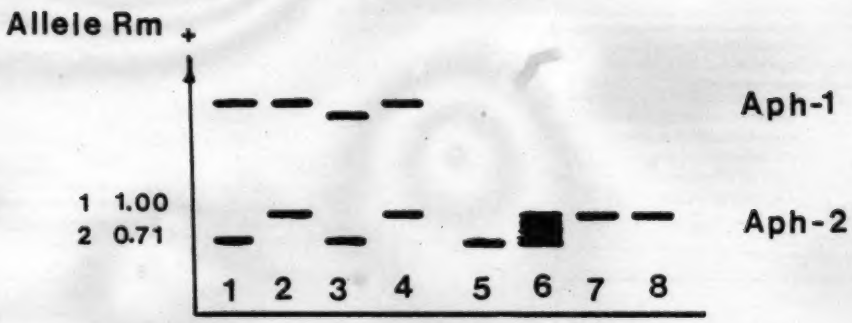

Figure 2.2. Diagram of acid phosphatase gels indicating haploid megagametophytes in lanes 1 to 4 and diploid embryo phenotypes in lanes 5 to 8. Lane: 5 = 22, 6 = 12, 7,8 = 11

Table 2.4. Observed segregation of allele 1 and 2 of enzyme Aph-2 by female, crown strata and year of collection.

Gamete Counts

		CROWN STRATA							
		Low		Medium		High		Total	
Female	Year	1	2	1	2	1	2	1	2
2	1975	7	3	9	1	4	6	20	10
	1976	6	3	5	4	6	4	17	11
	1977	5	4	7	3	7	3	19	10
	1978	missing		4	3	7	3	11	6
	Total	18	10	25	11	24	16	67	37

significantly deviant segregation ratio was observed in the crown strata sampled (M-1975, G_1^2=7.36; Table 2.5). However, the deviation of the pooled distribution was attributable to a consistent deficiency of the second allele at this locus. That is, while the distribution at each position sampled (except for M-1975) fit the expected distribution under the null hypothesis, there was an overall deficiency of the number 2 allele when all positions were pooled. This consistency is supported by the lack of detectable heterogeneity within any of the segregation summaries.

2.5 Malate dehydrogenase (Mdh) E.C. 1.1.1.37

Two zones of activity were observed on gels scored for Mdh (Figure 2.3). While there was evidence of four loci in this system, under the electrophoretic conditions for this study, only the two most cathodal loci could be scored reliably in the progeny. Both zones, Mdh-3 and Mdh-4, segregated in megagametophytes of heterozygous females for single-banded variants. Two alleles were observed at each of these loci. In both zones, embryos were either single-banded at the same position as the megagametophyte or triple-banded with one or two bands coincident with one or both bands observed in the megagametophytes. An intermediate band of approximately double the intensity was observed. This strongly suggests that both Mdh enzymes are functionally dimeric. Dimeric forms of Mdh have been reported in pitch

Table 2.5. Summary of the segregation analysis for Aph-2 by female, crown strata and year of collection.

	Distribution of Positions
X^2 Class	female 2
1	
2	L77 M76 H75,76
3	L76
4	L75 M77,78 H77,78
5	
6	
7	M75

Note, the Low strata in 1978 is missing.

Class intervals:

1=0-0.1015, 2=0.1016-0.455, 3=0.456-1.323, 4=1.324-2.706, 5=2.707-3.841 6=3.842-6.635, 7=>6.635

Adapted from Snedecor and Cochran (1973)

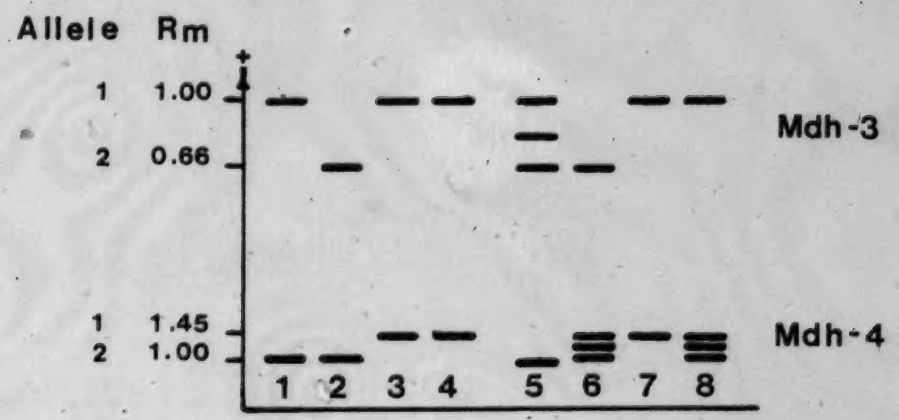

Figure 2.3. Diagram of malate dehydrogenase gels indicating haploid megagametophytes in lanes 1 to 4 and diploid embryo phenotypes in lanes 5 to 8. Mdh-3 Lane: 5 = 12, 6 = 22, 7,8 = 11. Mdh-4 Lane: 5 = 22, 6 = 12, 7 = 11, 8 = 12.

pine (*Pinus rigida* Mill.) by Guries and Ledig (1978).

2.5.1 Mdh-3

Overall segregation for the three heterozygous females at this locus fits the Mendelian expectation of a 1:1 ratio of gametes well (G_1^2=0.003; Table 2.6). There was, however, significant heterogeneity among females (G_3^2=9.98).

For tree number 10, the observed distribution of gametes deviated significantly from the expected distribution under the null hypothesis (G_1^2=6.89). This deviation was in the direction of an excess of the number 2 allele in the low crown stratum in 1977 and 1978, middle crown stratum in 1978 and high crown stratum in 1976 (Table 2.7). If the null hypothesis of 1:1 segregation of gametes in heterozygous females is true, we would expect approximately one chance deviation from the null hypothesis per female at the 5% level of probability. However, there were four positions of this female which deviated and all were in the same direction. Thus, it would seem that this female is not segregating in a 1:1 ratio spatially or temporally. For trees number 5 and 27, if we eliminate the mid-crown position in 1978 (only four gametes were tested for number 5), both are segregating temporally and spatially as expected under the null hypothesis of a 1:1 segregation ratio.

While the pooled distribution over both years and crown strata fit the expectations under the null hypothesis, there

Table 2.6. Observed segregation of allele 1 and 2 of enzyme Mdh-3, by female, crown strata and year of collection.

		Gamete Counts							
		CROWN STRATA							
		Low		Medium		High		Total	
Female	Year	1	2	1	2	1	2	1	2
5	1975	4	4	4	5	6	4	14	13
	1976	4	6	8	2	5	4	17	12
	1977	5	5	4	3	3	5	12	13
	1978	4	4	4	0	5	3	13	7
	Total	17	19	20	10	19	16	56	45
10	1975	6	2	2	7	5	5	13	14
	1976	7	3	3	4	2	8	12	15
	1977	2	8	6	4	3	7	11	19
	1978	0	7	1	6	3	6	4	19
	Total	15	20	12	21	13	26	40	67
27	1975	4	6	5	5	6	4	15	15
	1976	5	4	7	3	6	4	18	11
	1977	6	4	5	5	6	4	17	13
	1978	7	3	7	3	3	7	17	13
	Total	22	17	24	16	21	19	67	52
Summary	1975	14	12	11	17	17	13	42	42
	1976	16	13	18	9	13	16	47	38
	1977	13	17	15	12	12	16	40	45
	1978	11	14	12	9	11	16	34	39
	Total	54	56	56	47	53	61	163	164

Table 2.7. Summary of the segregation analysis of Mdh-3 by female, crown strata and year.

	Distribution of Positions		
X^2 Class	female 5	female 10	female 27
1	L75,77,78	H75	M75,77
2	L76 M75,M77 H75,76	M76,77	L75,76,77 H75,76,77
3	H77,78		
4		L75,76 H77	L75 M76,78 H78
5		M75	
6	M76,78	L77 M78 H76	
7		L78	

Class intervals:

1=0-0.1015, 2=0.1016-0.455, 3=0.456-1.323, 4=1.324-2.706, 5=2.707-3.841 6=3.842-6.635, 7=>6.635

Adapted from Snedecor and Cochran (1973)

was significant (G_3^2= 6.70) heterogeneity among trees in the middle crown stratum when pooled over years.

2.5.2 Mdh-4

The pooled segregation ratio for the four trees (numbers 4, 7, 11 and 29) heterozygous at Mdh-4 fits the expectations of a 1:1 segregation ratio (G_1^2=0.84), with no detectable heterogeneity (G_3^2=3.60; Table 2.8). Trees number 7 and 29 both exhibited one significant deviation from the null hypothesis, which was not unexpected (Table 2.9). However, trees number 4 and 11 had three and four significant comparisons, respectively. There was no consistent trend for the direction of the deviation among the three significant comparisons for tree 4. For tree 11, there was a tendency for a deficiency of the number 2 allele (in M-1978, H-1977 and H-1978), as well as significant heterogeneity of the segregation ratio of the middle crown stratum across the years (G_3^2=8.92). While no heterogeneity was detected in the high crown stratum, for all heterozygous trees there was an overall lack of fit to the expected distribution under the null hypothesis (G_1^2=8.40). This lack of fit was due to a deficiency of the 2 allele. Most of the deviant segregation ratios for Mdh-4 tended to occur in years 1977 and 1978. The pooled distribution for crown strata had significant heterogeneity among the three strata (G_3^2=7.18). The pooled distribution for the high crown strata did not fit the 1:1 segregation ratio (G_1^2=7.47). This lack

Table 2.8. Observed segregation of allele 1 and 2 of enzyme Mdh-4 by female, crown strata and year of collection.

		Gamete Counts							
		CROWN STRATA							
		Low		Medium		High		Total	
Female	Year	1	2	1	2	1	2	1	2
4	1975	9	10	5	5	5	3	19	18
	1976	11	9	6	4	8	2	25	15
	1977	4	6	2	8	3	7	9	21
	1978	9	1	5	4	2	3	16	8
	Total	33	26	18	21	18	15	69	62
7	1975	3	6	3	7	7	3	13	16
	1976	6	2	3	7	7	3	16	12
	1977	1	9	4	6	5	5	10	20
	1978	6	3	6	4	3	4	15	11
	Total	16	20	16	24	22	15	54	59
11	1975	6	4	4	6	6	4	16	14
	1976	3	4	8	5	5	5	16	14
	1977	6	4	2	8	9	1	17	13
	1978	5	5	6	1	9	1	20	7
	Total	20	17	20	20	29	11	69	48
29	1975	5	5	6	2	7	3	18	10
	1976	4	5	3	7	7	3	14	15
	1977	4	6	5	5	4	4	13	15
	1978	1	9	6	3	3	6	10	18
	Total	14	25	20	17	21	16	55	58
Summary	1975	23	25	18	20	25	13	66	58
	1976	24	20	20	23	27	13	71	56
	1977	15	25	13	27	21	17	49	69
	1978	21	18	23	12	17	14	61	44
	Total	83	88	74	82	90	57	247	227

Table 2.9. Summary of the segregation analysis of Mdh-4 by female, crown strata and year of collection.

	Distribution of Positions			
X^2 Class	female 4	female 7	female 11	female 29
1	L75 M75		L78 H76	L75 M77,H77
2	L76 M76,78 H78	M77,78 H77,78	L75,76,77 M75,H75	L76,77
3	H75	L75,78	M76	M78 H78
4	H77	L76 M75,76 H75,76		M75,76 H75,76
5				
6	M77 H76		M77,78	
7	L78	L77	H77,78	L78

Class intervals:

1=0-0.1015, 2=0.1016-0.455, 3=0.456-1.323, 4=1.324-2.706, 5=2.707-3.841 6=3.842-6.635, 7=>6.635

Adapted from Snedecor and Cochran (1973)

of fit was due to a deficiency of the number 2 allele.

2.6 6-Phosphogluconic dehydrogenase (6Pgd) E.C. 1.1.1.44

Two zones of activity were observed on gels scored for 6Pgd (Figure 2.4). While both zones were variable, only zone 2 provided consistent resolution in diploid embryos. Thus, this discussion will center only on variants detected in zone two (designated 6Pgd-2). Megagametophytes exhibited segregation in zone two of trees heterozygous for single-banded variants. Three alleles were observed in this zone. Embryos were either single-banded at the same position as the megagametophyte or triple-banded, with the alternate band coincident with the variants observed in the megagametophytes and an intermediate band of approximately double intensity. These results strongly suggest that the enzyme 6Pgd is functionally dimeric, as has been reported in loblolly pine (*Pinus taeda* L.) and pitch pine (Adams and Joly 1980a, and Guries and Ledig 1978).

2.6.1 6Pgd-2, allele 2

Overall segregation of the six trees heterozygous at this gene locus fits the expected distribution of gametes conditional on the null hypothesis of a 1:1 segregation ratio ($G_{\frac{1}{2}}$=0.43; Table 2.10) There was no significantly detectable heterogeneity among the pooled distributions of the heterozygous trees ($G_{\frac{1}{2}}$=0.49).

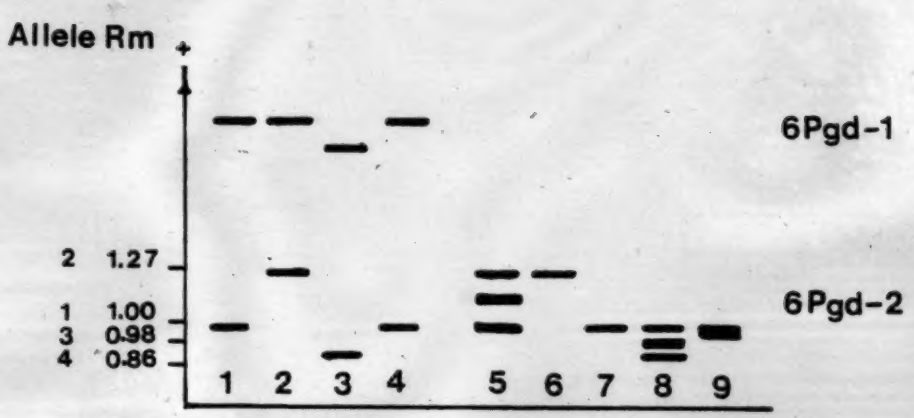

Figure 2.4. Diagram of 6-phosphogluconic dehydrogenase gel indicating haploid megagametopytes in lanes 1 to 4 and diploid embryo phenotypes in lanes 5 to 9. Lane: 5 =12, 6 = 22, 7 = 11, 8 = 14, 9 = 13.

Table 2.10. Observed segregation of allele 1 and 2 of enzyme 6Pgd-2 by female, crown strata and year of collection.

		Gamete Counts							
		CROWN STRATA							
		Low		Medium		High		Total	
Female	Year	1	2	1	2	1	2	1	2
1	1975	7	3	6	4	4	6	17	13
	1976	5	5	6	4	6	4	17	13
	1977	7	3	6	4	6	4	19	11
	1978	2	4	6	4	3	3	11	11
	Total	21	15	24	16	19	17	64	48
6	1975	2	8	5	4	1	9	8	21
	1976	4	4	5	3	7	3	16	10
	1977	2	4	4	3	3	7	9	14
	1978	7	3	4	4	7	3	18	10
	Total	15	19	18	14	18	22	51	55
9	1975	4	6	6	9	5	5	15	20
	1976	6	3	5	5	5	3	16	11
	1977	6	4	5	2	6	4	17	10
	1978	6	4	8	2	5	4	19	10
	Total	22	17	24	18	21	16	67	51
14	1975	3	7	missing		5	2	8	9
	1976	6	4	missing		6	4	12	8
	1977	5	5	3	6	8	2	16	13
	1978	8	2	2	8	4	6	14	16
	Total	22	18	5	14	23	14	50	46
18	1975	2	8	5	4	5	3	12	15
	1976	6	4	12	8	6	4	24	16
	1977	5	5	3	7	4	6	12	18
	1978	4	5	3	4	2	8	9	17
	Total	17	22	23	23	17	21	57	66
29	1975	5	5	4	4	6	4	15	13
	1976	2	7	5	5	6	4	13	16
	1977	5	5	4	6	7	1	16	12
	1978	3	7	5	4	3	6	11	17
	Total	15	24	18	19	22	15	55	58
Summary	1975	23	37	26	25	26	29	75	91
	1976	29	27	33	25	36	22	98	74
	1977	30	26	25	28	34	24	89	78
	1978	30	25	28	26	24	30	82	81
	Total	112	115	112	104	120	105	344	324

Trees number 1, 9 and 29 showed gamete distributions that fit the expected distributions (Table 2.11). There was no significant heterogeneity among the observed distributions for the crown strata or years sampled for these trees.

Tree 6 exhibited two distorted segregation ratios in the low and high crown strata in 1975. Both of these were due to a deficiency of the number 1 allele. The pooled year-distributions for this female had significant heterogeneity (G_H^2=10.70). The observed 1975 distribution deviated from expected (G_H^2=6.04) when pooled over crown strata. This lack of fit was attributable to a deficiency of the 1 allele.

Tree number 14 had three deviant segregation ratios. There was no consistent trend in the direction of deviation within years. However, L-1978 and H-1977 both had deficiencies of the 2 allele. When pooled over years, the observed gamete distribution in the middle crown stratum did not fit the expected 1:1 distribution (G_H^2=4.43). This was attributable to a deficiency of the 1 allele. As well, there was significant heterogeneity among crown strata (G_H^2=6.88).

Tree number 18 had two significant segregation distortion in L-1975 and H-1978. The lack of fit for both of these was due to a deficiency of the 1 allele. There was no heterogeneity or lack of fit for the year and crown-strata pooled distributions.

Table 2.11. Summary of the class membership for the segregation analysis of GPG-2, allele 2, by position, year and female.

	Distribution of Positions					
X^2 Class	female 1	female 6	female 9	female 14	female 18	female 29
1	L76,H78	L76,M78	M76,H75	L77	L77	L75,77,M75,76
2	M75,77,77,78 H75,76,77	M75,77	L75,77,78 H77,78	L76,H76,78	L76,78,M75,78 H76,77	M77,78,H75,76
3	L78	L77,M76	L76,M75,H76	M77	M76,H75	H78
4	L75,77	L78 H76,77 H78	M77	L75 H75	M77	L78
5						L76
6		L75	M78	L78,M78,H77	L75 H78	H77
7		H75				

Class intervals:

1=0-0.1015, 2=0.1016-0.455, 3=0.456-1.323, 4=1.324-2.706, 5=2.707-3.841

6=3.842-6.635, 7=>6.635

Adapted from Snedecor and Cochran (1973)

When all trees were considered, there was a tendency for aberrant segregation ratios to occur in 1977 and 1978. This trend was essentially independent of crown strata, although the low and high crown strata deviated more often than did the middle crown stratum.

2.6.2 6Pgd-2, allele 4

Only one female (number 7) among the 30 tested was heterozygous for alleles number 1 and 4. There were no significantly deviant segregation ratios from the expected distribution under the null hypothesis (Table 2.12). As well, there was no detectable heterogeneity among crown strata or years sampled for this female (Table 2.13)

2.7 Glucose-6-phosphate dehydrogenase (G6p) E.C. 1.1.1.49

One main zone of activity was observed on gels stained for G6p activity. There was, however, an ontogenetic modification of this zone of activity. After day four of germination, apparent activity in this zone progressively decreased and the apparent activity of another zone (anodal) increased at approximately the same rate. By approximately day ten, most detectable activity in the scored zone had disappeared and the main region of activity had shifted to the more anodal zone.

In trees apparently heterozygous for G6p, the megagametophytes showed segregation for single-banded

Table 2.12. Observed segregation of allele 1 and 4 of enzyme 6Pgd-2 by female, crown strata and year of collection.

Gamete Counts

		CROWN STRATA							
		Low		Medium		High		Total	
Female	Year	1	4	1	4	1	4	1	4
7	1975	3	6	9	6	6	4	18	16
	1976	5	3	7	8	7	3	19	14
	1977	5	5	8	5	10	10	23	20
	1978	7	2	6	8	9	7	22	17
	Total	20	16	30	27	32	24	82	67

Table 2.13. Summary of the segregation analysis for 6Pgd-2, allele 4 by female, position, and year of collection.

	Distribution of Positions
X^2 Class	female 7
1	L77 M76 H77
2	M78 H75,78
3	L75,76 M75,77
4	H76
5	L78
6	
7	

Class intervals:

1=0-0.1015, 2=0.1016-0.455, 3=0.456-1.323, 4=1.324-2.706, 5=2.707-3.841 6=3.842-6.635, 7=>6.635

Adapted from Snedecor and Cochran (1973)

variants. Embryos were either single-banded in the same position as the megagametophyte or exhibited one of two phenotypes (Figure 2.5). The embryos heterozygous for alleles 1 and 2 (= "12" embryos) commonly occurred with unresolved bands with the top and bottom of the zone of activity coincident with the relative migration distances observed for the two alleles in the megagametophytes. When clear banding patterns occurred they revealed a triple-banded phenotype in heterozygous embryos suggesting that the enzyme G6p is functionally dimeric. The other embryo phenotype, occasionally observed, appeared as an elongated, unresolved zone of weak activity extending to the origin. There was no evidence that the bands had migrated in a cathodal direction. No instances of clear banding patterns were observed for this embryo phenotype. Also, none of these phenotypes was observed in heterozygous females. Therefore, this phenotype was designated as "13".

The pooled distributions for both females heterozygous at G6p did not fit the expected 1:1 segregation ratio (G_1^2 =27.93; Table 2.14). As well, each female (numbers 15 and 27) exhibited segregation patterns significantly different from those expected under the null hypothesis (G_1^2=14.17 and 13.83 respectively). There was no heterogeneity between these two females for their observed segregation ratio. (G_1^2=0.07). In both cases, the lack of fit was attributable to a deficiency of the 2 allele. This deficiency was primarily attributable to samples M-1977 and L-1975 for

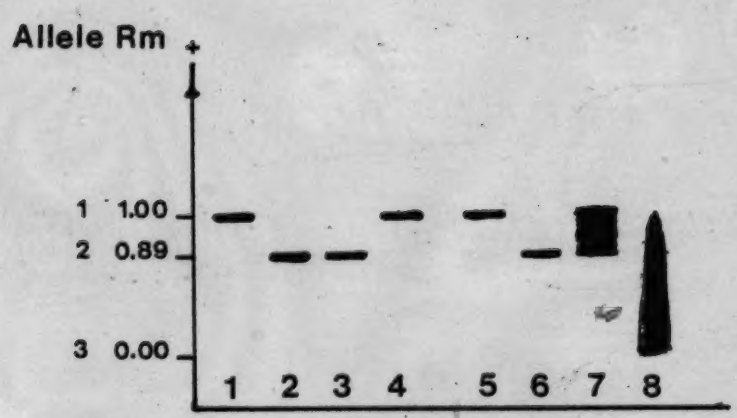

Figure 2.5. Diagram of glucose-6-phosphate dehydrogenase gel indicating haploid megagametophytes in lanes 1 to 4 and diploid embryo phenotypes in lanes 5 to 8. Lane: 5 = 11, 6 = 22, 7 = 12, 8 = 13.

Table 2.14. Observed segregation of allele 1 and 2 of enzyme G6p by female, crown strata and year of collection.

		Gamete Counts							
		CROWN STRATA							
		Low		Medium		High		Total	
Female	Year	1	2	1	2	1	2	1	2
15	1975	10	0	6	2	7	2	23	4
	1976	6	4	6	3	6	4	18	11
	1977	4	3	8	2	4	5	16	10
	1978	3	3	6	3	2	0	11	6
	Total	23	10	26	10	19	11	68	31
27	1975	5	5	7	3	9	1	21	9
	1976	5	5	6	3	8	2	19	10
	1977	5	5	5	5	7	3	17	13
	1978	7	4	5	3	10	0	22	7
	Total	22	19	23	14	34	6	79	39
Summary	1975	15	5	13	5	16	3	44	13
	1976	11	9	12	6	14	6	37	21
	1977	9	8	13	7	11	8	33	23
	1978	10	7	11	6	12	0	33	13
	Total	45	29	49	24	53	17	147	70

female 15, and H-1975, H-1976 and H-1978 for female 27 (Table 2.15) There was only one instance in the 23 observations where the number of the 2 alleles exceeded the number of 1 alleles observed (tree 15 H-1977). Therefore, we must conclude that these females cannot be considered as segregating in a 1:1 fashion for this locus.

2.8 Additional Variable Loci

Variation at a number of other loci also was detected only in embryos. The variants detected at each locus are briefly described below. At each of these other loci, no illegitimate progeny (i.e., progeny homozygous for an allele not in the mother tree) were obtained. Also, no progeny were detected that were heterozygous for alleles not contained in the maternal genome. In most of these cases, the megagametophyte was coincident in relative migration to the hypothesized allele derived from the maternal component contained in the embryo. As well as the loci reported below, Isocitrate dehydrogenase (E.C. 1.1.1.42 Idh) was scored. However, no variation was detected at this locus.

2.8.1 Aspartate aminotransferase (Aat-1) E.C. 2.6.1.1

A total of four alleles were detected in the progeny at Aat-1. In all cases, progeny were triple banded, suggesting that the enzyme Aat is functionally dimeric (Figure 2.6), as has been reported in loblolly pine (Adams and Joly 1980a).

Table 2.15. Summary of the segregation analysis for G6p by female, crown strata and year of collection.

	Distribution of Positions	
X Class	female 15	female 27
1	L78	L75,76,77 M77
2	L76,77 H76,77	
3	M76,78	L78 M76,78
4	M75	M75,H77
5	H75	
6	M77	H76
7	L75	H75,78

Class intervals:

1=0-0.1015, 2=0.1016-0.455, 3=0.456-1.323, 4=1.324-2.706, 5=2.707-3.841 6=3.842-6.635, 7=>6.635

Adapted from Snedecor and Cochran (1973)

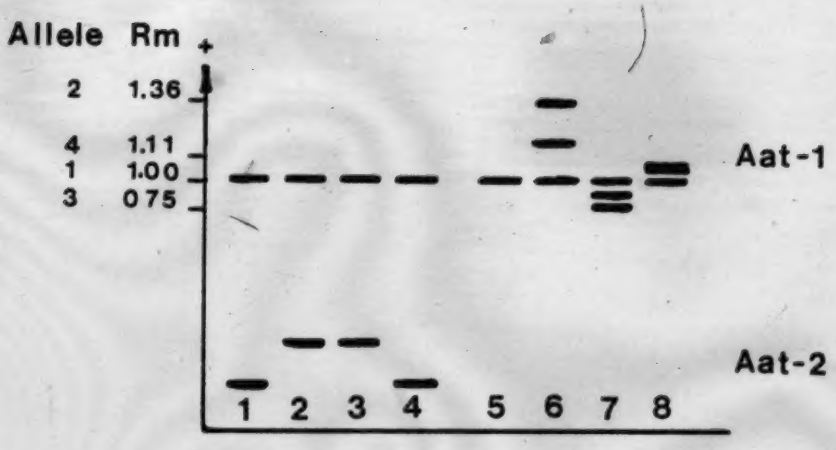

Figure 2.6. Diagram of aspartate aminotransferase gel indicating haploid megagametophytes in lanes 1 to 4 and diploid embryo phenotypes in lanes 5 to 8. Lane: 5 = 11, 6 = 12, 7 = 13, 8 = 14.

2.8.2 Alcohol dehydrogenase (Adh) E.C. 1.1.1.1

Two alleles were detected in the progeny at Adh. All progeny were triple-banded in the heterozygous condition, suggesting that the enzyme Adh is functionally dimeric (Figure 2.7).

2.8.3 Glutamate dehydrogenase (Gdh) E.C. 1.4.1.2

Only one additional allozyme variant was detected in Gdh. In embryos heterozygous at Gdh, the top of the bands did not extend to the same relative position as did the haploid megagametophyte (Figure 2.7). This condition has been observed and reported in loblolly pine (Adams and Joly 1980a). Since the enzyme Gdh appears to be a multimer, it is not expected that the bands should be coincident in their migration (Pryor 1974).

2.8.4 Malic enzyme (Me) E.C. 1.1.1.40

A total of three alleles were detected in the progeny of Me. In heterozygous condition, embryos were double-banded, suggesting that the enzyme Me is a monomer (Figure 2.8).

2.9 Linkage Analaysis

Two trees (numbers 7 and 27) were multiply heterozygous for Mdh-4 and 6Pgd-2 (allele 4), and Mdh-3 and G6p, respectively. The G^2 analysis of 1:1:1:1 distribution of

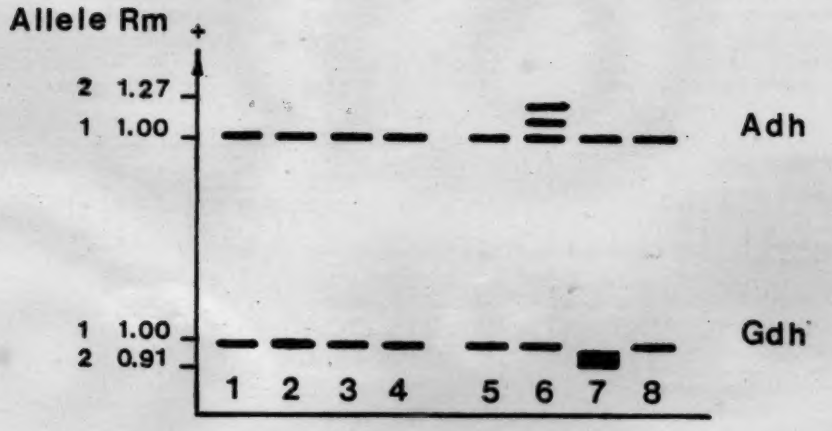

Figure 2.7. Diagram of alcohol dehydrogenase and glutamate dehydrogenase gel indicating haploid megagametophytes in lanes 1 to 4 and diploid embryo phenotypes in lanes 5 to 8. Adh Lane: 5 = 11, 6 = 12, 7 = 11, 8 = 11. Gdh Lane: 5 = 11, 6 = 11, 7 = 12, 8 = 11.

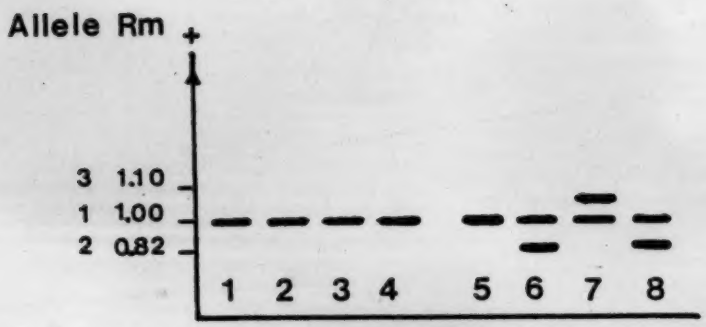

Figure 2.8. Diagram of malic enzyme gel indicating haploid megagametophytes in lanes 1 to 4 and diploid embryo phenotypes in lanes 5 to 8. Lane: 5 = 11, 6 = 12, 7 = 13, 8 = 12.

gametes revealed no evidence of linkage between the loci of either pair.

2.10 Discussion

Sampling theories developed to detect a heterozygous female by assaying gametes are based on the seemingly reasonable assumption of a 1:1 segregation ratio of gametes in heterozygous females (Morris and Speith 1978). This assumption likely is valid only at meiosis. However, studies of conifer megagametophytes do not assess the segregation ratio at meiosis, but at a time considerably later. There are a number of events which must occur before a filled seed, suitable for genotyping, matures. These include both pre- and post-zygotic events. Pre-zygotic events include the absence of meiotic irregularities, pollination, and fertilization of the particular ovule. At this latitude, meiosis and the mature seed in *Pinus* are separated by an interval of approximately 17 months (Owens *et al.* 1981). Meiotic irregularities are known to occur in all species of *Pinus* studied (Saylor and Smith 1966). These irregularities may take the form of:

1. precocious disjunction, which leads to the production of univalents and
2. failure of chiasma terminalization which results in lagging chromosomes as well as chromosome breakage and fragments.

In addition, paracentric inversions were observed in virtually all species studied.

Post-zygotic factors necessary to ensure development of the embryo include a host of environmental parameters, general nutritional status of the female and a number of genetic factors not completely understood. A breakdown in the post-zygotic phase normally results in abortion of the embryo and desiccation of the female gametophytic tissue (Orr-Ewing 1957). The net result is a normal appearing, but empty seed. Additional constraints, such as germination of the seed with a 2-3 mm radicle, impose further obstacles between meiosis and the assayed megagametophyte. The sampling strategy of Morris and Speith (1978) only indicates the probability of seeing a heterozygote; it does not assess the segregation ratio of a heterozygous female.

What factors could be responsible for observed segregation distortion? While it is impossible to test conclusively a given hypothesis with these data, some of the more obvious possible factors responsible can be mentioned. Paracentric inversions produce dicentric and acentric chromosomes. However, this should simply increase the rate of pollen abortion and proportion of non-functional megagametophytes. Thus, apart from a reduction in fertility, they should result in no segregation distortion. Given that they occur at random, the other two meiotic irregularites noted also should not cause segregation disortions. Thus, meiotic irregularities by themselves do not appear to be

able to adequately account for segregation distortion. Meiotic disturbances such as segregration preference or meiotic drive mechanisms are known to occur in plants, but not in gymnosperms (Rhoades 1952). Furthermore, they would not account for the observed heterogeneity among females or among sampling positions within a female.

One of the more obvious underlying explanations for the observed segregation distortion would be the genetic composition of the pollen pool. The pollen pool must be considered in most of these discussions since maternal segregation was indirectly assayed after fertilization. Conifers typically have a high embryonic-lethal genetic load (Fowler 1965, Koski 1971, Sorensen 1969). In Douglas-fir, the mean number of lethal equivalents is nine to 10 per zygote (Sorensen 1969). Estimates for other species range from one to two per zygote in *Drosophila pseudoobscura* (Dobzhansky *et al.* 1963) and *Tribolium* (Levene *et al.* 1965) to three to five per zygote in man (Morton *et al.* 1965). In forest trees, the majority of this load is expressed in early embryonic development, particularly after self-fertilization, although some is continually expressed from seed germination through all stages to maturity (Orr-Ewing 1957, 1965). However, it is also conceivable that embryonic lethals could be expressed upon mating with related individuals (e.g., parent-offspring, first cousins, etc.), as well as with random matings. Furthermore, self-fertilization in conifers more often results in a

higher proportion of empty seeds than cross-fertilization (Hadders and Koski 1975, Rudolph 1976), although a large proportion of empty seeds also is observed under normal field conditions (Anderson 1965). After controlled self-pollination, the progress of fertilization and subsequent embryogenesis generally proceed normally until shortly after the suspensor stage (i.e., the begining of the embryo proper) in *Picea abies* and *Pinus sylvestris* L. (Koski 1971). Orr-Ewing (1957) showed that development of the embryo following self-fertilization continued normally until two or three weeks after fertilization in *Pseudotsuga menziesii* (Mirb.) Franco. Given the results presented here, one explanation for the aberrant segregation ratios observed would be either linkage of the isozyme marker with some embryonic lethal or direct lethality of the isozyme itself. However, since both types of homozygous progeny were observed, lethality of the isozyme would not be complete. Sorensen (1967) has formulated some models to accomodate aberrant segregation ratios observed with quantitative characters.

In situations where there was no heterogeneity among females (or sampling positions within a female), we could hypothesize that either the pollen pool was homogeneous (for outcrossing) or that the genetic background of these females did not differ for lethals or semi-lethals upon selfing.

When heterogeneity among females or among sampling positions within a female is observed, it would seem

reasonable to infer that the pollen pool is heterogeneous in space and time. Alternatively for inter-female differences, different genetic backgrounds could be responsible for the observed departure from 1:1 segregation ratios. Homogeneity and heterogeneity of the pollen pool also refer to the proportion of embryos derived as a direct result of auto-fertilization, as well as strictly genetic changes in the external pollen pool. From these experiments and analyses, it is not possible to unambiguously identify an embryo resulting from cross fertilization with a pollen allele identical to the allele of the mother (identical in type or homozygous) or from an auto-fertilization event (identical in state or autozygous).

This is not the only report of distored segregation ratios observed in electrophoretic investigations of coniferous populations. Adams and Joly (1980a) reported segregation distortion in a seed orchard of *Pinus taeda*. Eckert *et al.* (1981) observed aberrant segregation ratios in a clonal seed orchard of *Pinus strobus* L. Lundkvist (1974) indicated distorted segregation ratios for natural populations of *Picea abies*. In their study of natural populations of *Pinus ponderosa* Laws., O'Malley *et al.* (1979) observed segregation disorders at several gene loci. Rudin (1975, 1977) and Rudin and Ekberg (1978) reported segregation distortion from the expected 1:1 ratios in different natural populations of *Pinus sylvestris*. Thus, segregation distortions from 1:1 expectations seem to be the

rule, rather than the exception in coniferous populations. These segregation distortions are most likely the result of linkage between the electrophoretic marker and an embryonic lethal or semi-lethal gene.

These types of segregation distortion should be considered when estimates of population allele frequency are made from bulked samples of megagametophytes, particularly when a small number of individuals are represented in the sample. Furthermore, these aberrant segregation ratios affect the binominal probability statement about detection of heterozygosity of an individual plant. Thus, it may be necessary to include one or two additional megagametophytes to compensate for these sources of bias.

2.11 Literature Cited

Adams, W.T., and R.J. Joly 1980a. Genetics of allozyme variants in loblolly pine. J. Hered. 71:33-40.

Adams, W.T., and R.J. Joly 1980b. Linkage relationships among twelve allozyme loci in loblolly pine. J. Hered. 71:199-202.

Allard, R.W. 1975. The mating system and microevolution. Genetics 79 (suppl):115-126.

Anderson, E. 1965. Cone and seed studies in Norway spruce (*Picea abies* (L.) Karst.). Studia Forestalia Suecica 23:1-214.

Bartels, H. 1971. Genetic control of multiple esterases from

needles and macrogametophytes of *Picea abies*. Planta 99:283-289.

Bergmann, F. 1973. Genetic studies in *Picea abies* with the aid of isozyme identification. II. Genetic control of esterase and leucine aminopeptidase isozymes in haploid endosperm of dormant seeds. Theor. Appl. Genet. 43:222-225 (German with English summary).

Bergmann, F. 1974. The genetics of some isoenzyme systems in spruce endosperm (*Picea abies*). Genetica 6:353-360.

Brown, A.H.D. 1979. Enzyme polymorphisms in plant populations. Theor. Pop. Biol. 15:1-42.

Clayton, J.W., and D.N. Tretiak. 1972. Amino-citrate buffers for pH control in starch gel electrophoresis. J. Fisheries Res. Board Can. 29:1169-1172.

Conkle, M.T. 1971. Inheritance of alcohol dehydrogenase and leucine aminopeptidase isozymes in knobcone pine. For. Sci. 17:190-194.

Dobzansky, T., B. Spaskey and T. Tidwell. 1963. Genetics of natural populations. XXXII. Inbreeding and the mutational and balanced genetic loads in natural populations of *Drosophila pseudoobscura*. Genetics 48:361-373.

Eckert, R.T., R.J. Joly and D.B. Neale. 1981. Genetics of isozyme variants and linkage relationships among allozyme loci in 35 eastern white pine clones. Can. J. For. Res. 11:573-579.

El-Kassaby, Y.A., F.C.H. Yeh and O. Sziklai. 1982.

Inheritance of allozyme variants in costal Douglas-fir (*Pseudotsuga menziesii* var. *menziesii*) Can. J. Genet. Cytol. 24:325-335.

El-Kassaby, Y.A., O. Sziklai and F.C.H. Yeh 1982b. Linkage relationships among 19 polymorphic allozyme loci in coastal Douglas-fir (*Pseudotsuga menziesii* var. *menziesii*). Can. J. Genet. Cytol. 24:101-108.

Fowler, D.P. 1965. Effects of inbreeding in red pine, *Pinus resinosa* Ait. IV. Comparison with other northeastern *Pinus* species. Silvae Genetica 14:76-81.

Guries, R.P., and F.T. Ledig. 1978. Inheritance of some polymorphic isozymes in pitch pine (*Pinus rigida* Mill.). Heredity 40:27-32.

Guries, R.P., S.T. Friedman and F.T. Ledig. 1978. A megagametophytic analysis of genetic linkage in pitch pine (*Pinus rigida* Mill). Heredity 40:309-314.

Hadders, G., and V. Koski. 1975. Probability of inbreeding in seed orchards. p 108-117 in Seed orchards (R. Faulkner ed.) For. Comm Bull. 54. H.M.S.O. London

Koski, V. 1971. Embryonic lethals of *Picea abies* and *Pinus sylvestris*. Comm. Inst. For. Fenn. 75(3):1-30.

Levene, H., I. M. Lerner, A. Sokaloff, F. K. Ho and I. R. Franklin. 1965. Genetic load in *Tribolium* Proc. Nat. Acad. Sci. (USA) 53:1042-1050.

Lewontin, R.C. 1974. The genetic basis of evolutionary change. Columbia Univ. Press, N.Y.

Lundkvist, K. 1974. Inheritance of leucine aminopeptidase

isozymes in *Picea abies* K. (Koch.). Hereditas 76:91-96.

Lundkvist, K. 1975. Inheritance of acid phosphatase isozymes in *Picea abies*. Hereditas 85:67-74.

Lundkvist, K. 1977. Inheritance of esterases in needles and endosperms of Norway spruce (*Picea abies* K.). Hereditas 87:27

Lundkvist, K. 1979. Allozyme frequency distributions in four Swedish populations of Norway spruce (*Picea abies* K.). I. Estimates of genetic variation within and among populations, genetic linkage and mating system parameters. Hereditas 90:127-143.

Lundkvist, K., and D. Rudin. 1977. Genetic variation in 11 populations of *Picea abies* in Sweden as determined by isozyme analysis. Hereditas. 85:67-74.

Morris, R.W., and P.T. Speith. 1978. Sampling strategies for using female gametophytes to estimate heterozygosity in conifers. Theor. Appl. Genet. 51:217-222.

Morton, N.E., J.F. Crow and H.J. Muller. 1956. An estimate of the mutational damage in man from data on consanguineous marriages. Proc. Nat. Acad. Sci. (USA) 42:855-863.

Nevo, E. 1978. Genetic variation in natural populations: Patterns and theory. Theor. Pop. Biol. 13:121-177.

O'Malley, D.M., F.W. Allendorf and G.M. Blake. 1979. Inheritance of isozyme variation and heterozygosity in *Pinus ponderosa*. Biochem. Genet. 17:233-250.

Orr-Ewing, A.L. 1957. A cytological study of the effects of

self-pollination of *Pseduotsuga menziesii* (Mirb.) Franco. Silvae Genetica 6:179-185.

Orr-Ewing, A.L. 1965. Inbreeding and single crossing in Douglas-fir. For. Sci. 11:279-290.

Owens, J.N., S.J. Simpson and M. Molder. 1981. Sexual reproduction of *Pinus contorta*. I. Pollen development, the pollination mechanism, and early ovule development. Can. J. Bot. 59:1828-1843.

Pryor, A.J. 1974. Allelic glutamic dehydrogenase isozymes in maize -- a single hybrid isozyme in heterozygotes? Heredity 32:397-401.

Rick, C.M., J.R. Forbes and M. Holle. 1977. Genetic variation in *Lycopersicon pimpinellifolium*: Evidence of evolutionary change in mating systems. Plant. Syst. Evol. 127:139-170.

Ridgeway, G.J., S.W. Sherburne and R.D. Lewis. 1970. Polymorphisms in the esterases of Atlantic herring. Trans. Am. Fisheries Sco. 99:147-151.

Rhoades, M.M. 1952. Preferential segregation in maize. p. 60-80 in Heterosis. Iowa State College Press, Ames.

Rudin, D. 1975. Inheritance of glutamate oxalate-transaminases (GOT) from needles and endosperms of *Pinus sylvestris*. Hereditas 80:296-300.

Rudin, D. 1977. Leucine aminopeptidases (LAP) from needles and endosperms of *Pinus sylvestris* L.-a study of inheritance and linkage of allozymes. Hereditas 85:219-226.

Rudin, D., and I. Ekberg. 1978. Linkage studies in *Pinus sylvestris* L. using macrogametophyte allozymes. Silvae Genetica 27:1-12.

Rudolph, T.D. 1976. Cone set, seed yield, seed quality, and early seedling development of S_2 generation jack pine. p 42-60 in Proc. 10th Central States Forest Tree Improv. Conf. Purdue Univ.

Saylor, L. C., and B. W. Smith. 1966. Meiotic irregularity in species and interspecific hybrids of Pinus. Amer. J. Bot. 53:453 -468.

Siciliano, M.J., and C.R. Shaw. 1976. Separation and visualization of enzymes on gels. p 185-209 in Chromatographic and electrophoretic techniques vol. 2 (I. Smith ed.). Heinemann, London.

Shaw, D.V., A.L. Kahler and R.W. Allard. 1981. A multilocus estimator of mating system parameters in plant populations. Proc. Nat. Acad. Sci. (USA) 78:1298-1302.

Sorensen, F.C. 1967. Linkage between marker genes and embryonic lethal factors may cause disturbed segregation ratios. Silvae Genetica 16:132-134.

Sorensen, F.C. 1969. Embryonic genetic load in coastal Douglas-fir *Psudotsuga menziesii* var. *menziesii*. Amer. Nat. 103:389-398.

Tobolski, J.J. 1979. Inheritance of polymorphic isozymes in jack pine (*Pinus banksiana* Lamb.). p 147-154 in Proc., First North Central Tree Improv. Conf.

Yeh, F.C.H., Y.A. El-Kassaby, M.A.K. Khalil and D.C.

Chipman. 1982. Inheritance and linkage relationships of allozyme variants in black spruce. (Submitted Can. J. Genet. Cytol.).

Yeh, F.C.H., and D. O'Malley. 1980. Enzyme variations in natural populations of Douglas-fir, *Pseudotsuga menziessi* (Mirb.) Franco, from British Columbia. 1. Genetic variation patterns in coastal populations. Silvae Genetica 29:83-92.

3. ESTIMATION OF MATING SYSTEM AND EFFECTIVE POPULATION SIZES

3.1 Introduction

In an infinitely large population, if all of the assumptions of the Hardy-Weinberg law are satisfied, gene frequencies and amounts of variation remain constant from generation to generation. However, in most populations, migration, mutation, selection, non-random mating and finite size operate to cause either directional or stochastic changes in the population from generation to generation. Non-random mating in the form of selfing or consanguineous mating does not result in changes of allele frequencies from one generation to the next (Crow and Kimura 1970). However, the genotypic distributions and homozygosity are modified as a function of the level of consanguineous matings. Other forms of non-random mating, such as positive assortative mating, cause an increase in the frequency and homozygosity of the target gene and others linked to it, but do not affect the rest of the genome (Crow and Felsenstein 1968).

The finite size of a population can have two effects on the structure of the population. The first is to increase the average degree of relationship among individuals from one generation to the next. This has been extensively studied from two approaches: correlation of uniting gametes (Wright 1965) and probability of identity by descent

(Malecot 1969). In addition, this finite size induces a fluctuation or variance associated with allele frequencies between successive generations. These situations, namely an increase in the average degree of relationship and variance in allele frequency, both act to reduce the actual number of individuals participating in mating from the total number of individuals in the population. From the inbreeding aspect, this occurs as a result of a reduced number of independent or unique genomes participating in the mating. The variance in allele frequency concept is derived from sampling theory, whereby a smaller number of samples from the population results in a larger variance associated with the mean. These two concepts of reduction in the genetic sample size of the population are termed the inbreeding effective number and the variance effective number, respectively (Crow 1954, Kimura and Crow 1963).

In this chapter, estimates are made of time and space variation in the mating system and variance effective population number in a stand of jack pine. As well, initial estimates are made of the variance effective population numbers for the male gametophytic population.

3.2 Materials and Methods

Collection of the open-pollinated seeds, tissue preparation, electrophoretic conditions and enzymes assayed were as previously indicated. Estimation of mating system

parameters and variance effective population number were made for temporal and spatial dimensions in the population. The time dimension was represented by the four pollination events of 1975, 1976, 1977 and 1978. The spatial dimensions were represented by three strata (Low, Middle and High) within the sexually active crown of a tree.

The maternal gamete contributed to each embryo was determined for each of the nine variable loci studied. Therefore, the observed pollen allele contributed to each embryo can also be determined. These data will be called the 'observed pollen pool' as opposed to the 'outcrossed pollen pool'. The latter component will be estimated from the mating system. Expected heterozygosity (h^2) was calculated for each variable locus in each population by the formula:

$$h^2 = 1 - \sum_i \sum_j p_{ij}^2 \qquad -1-$$

where p_{ij} represents the estimate of the frequency of the i th allele in the j th population. The observed genotypic distribution was compared to the expected genotypic distribution under Hardy-Weinberg equilibrium by means of a maximum likelihood G^2 (asymptotic X^2) analysis after an adjustment for small population size (Levene 1949). The proportional decrease in heterozygosity (F) relative to that expected under panmixia (i.e., Hardy-Weinberg equilibrium) was calculated for both the 30 mature trees and the temporally and spatially subdivided filial populations

according to the formula:

$$F = 1 - \frac{h^2}{H} \qquad -2-$$

where H is the observed proportion of heterozygotes and h^2 is as previously defined in 1. This quantity is known as Wright's fixation index.

The methods presented by Yasuda (1969) were used to estimate the variance effective population number (see Appendix 1 for details of the method). These estimates are based on the allele frequency estimates in the mature population and the observed pollen pool and compared to the estimates in the filial generation. If the sample of the mature population accurately represents the population allele frequencies no migration occurs, mating is at random, and all individuals are contributing equally to the pollen pool, then the mature population and observed pollen pool should give comparable estimates. As previously discussed, a constant amount of selfing does not change the allele frequencies between generations.

Estimation of single-locus mating system parameters (p(i), and s where p(i) is the frequency of the i th allele in the outcrossed pollen pool, i=1,..,k, s is the selfing rate and t=1-s is the outcrossing rate) were made using the methods of Brown *et al.* (1975) and Cheliak *et al.* (1982) (see Appendix 2 for details of the EM method). Estimation of mating system parameters using multi-locus data were made

using the methods of Shaw *et al.* (1981).

To assess the homogeneity of the pollen pool reaching each female, the number of homozygous and heterozygous embryos borne by a female was entered into a 2xf contingency table (Brown *et al.* 1975), where f is the number of homozygous females, and tested by means of a maximum likelihood G^2 analysis. A sum of squares simultaneous test procedure (STP; Sokal and Rohlf, 1969) was applied to delineate homogeneous subsets if the G^2 test of homogeneity was rejected.

3.3 Results

Five of the ten loci surveyed were found to be variable in the mature population. However, because of systematic deviations from expected Mendelian segregation, G6p has been excluded from all further calculations. Two alleles were observed in each of Aph-2, Mdh-3, and Mdh-4, and three alleles were observed in 6Pgd-2 (Table 3.1). Levels of heterozygosity ranged from a low of 0.0328 (Aph-2) to a high of 0.2094 (6Pgd-2), with a mean of 0.1052 (considering only the variable loci) or 0.0526 (considering all loci included in the study). The average number of alleles per locus was 1.67 for all loci and 2.2 for only the variable loci in the mature population. While this is a biased sample of loci, this estimate is low in comparison to most gymnosperms (Hamrick *et al.* 1979) and should not be construed as

Table 3.1. Allele frequencies, heterozygosities and test results for Hardy-Weinberg equilibrium distributions in the mature and filial populations.

		Locus			
Population	Allele	Aph-2	Mdh-3	Mdh-4	6Pgd-2
Mature	1	.983	.950	.933	.883
	2	.017	.050	.067	.100
	4	-	-	-	.017
	h^2	.033	.095	.124	.210
	†G^2	.000 (1)	.105 (1)	.214 (1)	.795 (3)
	F	-.016±.053	-.053±.026	-.072±.034	-.114±.087
	N	30	30	30	30
Filial	1	.978	.953	.948	.907
	2	.022	.047	.052	.083
	4	-	-	-	.010
	h^2	.043	.090	.099	.170
	G^2	.32(1)	16.08(1)	1.25(1)	17.75(3)
	F	-.012±.013	.079±.039	.015±.039	-.043±.017
	N	3397	3498	3496	3454

†Test for Hardy-Weinberg equilibrium, degrees of freedom shown in the brackets.

representative of this stand or the species. The distribution of genotypes in the mature population fits the expectations of Hardy-Weinberg equilibrium for all of the loci surveyed. However, as indicated by the negative F values, there was a consistent excess of heterozygotes in the mature population. The excess of heterozygotes over panmictic expectations was negatively related to the level of heterozygosity at the particular locus.

3.4 Filial population

3.4.1 Loci in Hardy-Weinberg equilibrium

No distortions from Hardy-Weinberg equilibrium were observed for any pooled filial distribution at Aat1, Adh, Aph-2, Gdh, Mdh-4 and Me. For those loci fixed in the mature population, there was a small, but consistent excess of observed heterozygotes in the filial population (Table 3.2). For Aph-2, there were two instances of non-significant excesses of homozygotes observed in the individual sampling positions. While the mean F value indicated an excess of observed heterozygotes at Aph-2, there was a large range in the individual sampling positions contributing to this mean (-0.0333 to 0.1333). A deficiency of heterozygotes was observed in the pooled distributions for 1976 and the high crown stratum for this locus. The mean F value for Mdh-4 indicated a small excess of observed homozygotes (0.0152),

Table 3.2. Summary of inbreeding coefficients (F) calculated from equation 2 for year and crown stratum subdivisions in the filial populations.

Locus	Stratum	1975	1976	1977	1978	Mean
Aph-2	Low	-.0112	-.0297	-.0270	0	-.0245
	Middle	-.0287	-.0288	-.0338	-.0239	-.0290
	High	-.0202	.1333	.0691	-.0198	.0222
	Mean	-.0222	.0158	-.0275	-.0207	-.0124
Mdh-3	Low	.1073	.1053	.0746	.0419	.0825
	Middle	.0872	.1323	.0334	.3333	.1292
	High	.0255	-.0528	.1763	.0037	.0341
	Mean	.0724	.0549	.0954	.0918	.0796
Mdh-4	Low	.0870	.0236	-.0756	-.0703	-.0183
	Middle	-.0407	-.0702	.0536	.2308	.0305
	High	-.0647	.3566	-.0478	-.0425	.0393
	Mean	.0076	.0764	-.0255	.0211	.0152
6Pgd-2	Low	-.0304	-.0909	-.0589	-.0945	-.0670
	Middle	.0888	-.0873	.0858	-.0522	.0052
	High	-.0956	-.0907	-.0853	.0013	-.0649
	Mean	-.0152	-.0896	-.0238	-.0457	-.0431

even though more than 50% of the observed progeny distributions indicated an excess of heterozygotes relative to that expected under panmixia. Two instances of significant excess of homozygotes were observed in the individual sampling positions (Middle 1978, High 1976).

3.4.2 Loci not in Hardy-Weinberg equilibrium

The pooled distributions for two loci, Mdh-3 and 6Pgd-2, were not in Hardy-Weinberg equilibrium. There were two instances of significant departures from the expected Hardy Weinberg equilibrium distribution for the individual sampling positions at Mdh-3 (Middle 1978, High 1977). The pooled year and crown strata summaries showed excesses of homozygotes, although only 1977 and 1978 and the low and middle strata had significant excesses. None of the individual positions showed significant departures from the expected distributions for 6Pgd-2. In the pooled year and crown strata summaries there was usually an excess of heterozygotes observed, although this excess was only significant for 1976.

3.4.3 Summary of levels of variation

Aph-2 typically showed an excess of variation in the filial populations as compared to the mature population. The increase in mean heterozygosity was approximately 33% when pooled over the twelve sampling positions. For Aph-2 there was an excess of heterozygotes, as measured by the F

statistics, in both the mature population (-.016) and the filial population (-.012). The remaining three variable loci in the mature population showed an increase in frequency of the common allele and a decrease in the expected level of variability from observations in the mature and filial populations. Although the observed progeny distribution for Mdh-3 did not fit panmictic expectations in terms of mean allele frequency of the variable loci studied, it was the most similar to the mature population. Overall, there was about a 6% decrease in the level of variability in the filial population at Mdh-3. However, there was a significant change in genotypic distributions from the mature to filial populations at Mdh-3, as measured from the F distributions. Both Mdh-4 and 6Pgd-2 showed approximately a 20% reduction in variability from the mature crop to the filial generation. However, the observed progeny distribution at Mdh-4 was consistent with panmictic expectations, while at 6Pgd-2 there was a significant excess of heterozygotes observed. In addition, the mature population at Mdh-4 showed an excess of heterozygotes, while there was a net deficiency of heterozyotes in the filial population. Excesses of heterozyotes were observed in both the mature and filial populations at 6Pgd-2. However, the excess was about 2.7 times greater in the mature than in the filial population. The change in the inbreeding value from the mature population to the filial populations was not significant.

3.5 Estimation of the mating system

3.5.1 Comparison of methods

Simultaneous maximum-likelihood estimates of the outcrossed pollen allele frequencies (p(i)) and the outcrossing rate (t) for Aph-2, Mdh-3, Mdh-4 and 6Pgd-2 were made using the methods of Brown *et al.* (1975) and Cheliak *et al.* (1982) (Table 3.3). In general there is good correspondence among the two methods in the maximum-likelihood estimates of the pollen pool allele frequencies, the exception being the middle crown stratum in 1977 for Aph-2, where the two estimates differed by 7%. The most striking difference between the two methods is in the estimates of outcrossing rates. While there is a strong rank correlation between outcrossing estimates derived from the two methods ($r=0.713$, 45 df $P<0.05$), more than half of the estimates from the estimation procedure of Brown *et al.* (1975) are greater than one and, thus outside of the natural biological parameter range. This 'excess of outcrossing' is related to the underlying observed genotypic distribution in the filial populations. Comparison of the observed fixation index (Table 3.2) in those of cases where $t > 1.0$ shows a complete correspondence between negative F values (i.e., excess of observed heterozygotes) and outcrossing rates greater than one. While the method of Brown *et al.* (1975) is insensitive to the genotypic structure of the mature population, excesses of heterozygotes in the filial

Table 3.3. Estimates of mating system parameters of jack pine using the methods of Brown et al. (1975) (=B) and Cheliak et al. (1982) (=EM).

		LOCUS																	
		Aph-2				Mdh-3				Mdh-4				6Pgd-2				Mean	
		B		EM		B		EM		B		EM		B		EM		B	EM
Crown Strata	Year	$\hat{P}_i$	$\hat{t}$	$\hat{P}_i$	$\hat{t}$	$\hat{P}_i$	$\hat{t}$	$\hat{P}_i$	$\hat{t}$	$\hat{P}_i$	$\hat{t}$	$\hat{P}_i$	$\hat{t}$	$\hat{P}_i$	$\hat{t}$	$\hat{P}_i$	$\hat{t}$	$\hat{t}$	$\hat{t}$
Low	75	.98	1.03	.98	.84	.95	.81	.95	.77	.97	.85	.97	.95	.94	1.01	.95	1.00		
	76	.96	1.09	.95	.87	.96	.80	.95	.73	.96	.97	.96	.96	.94	1.15	.94	1.00		
	77	.97	1.07	.96	.85	.95	.80	.95	.93	.97	1.07	.95	.96	.93	1.11	.92	.98		
	78	--	--	--	--	.98	.82	.96	.74	.99	1.03	.97	.87	.95	1.15	.93	1.00		
Mean			1.06		.85		.81		.79		.98		.89		1.11		1.00	0.99	0.88
Middle	75	.96	1.08	.95	.89	.95	.77	.95	.86	.99	1.02	.99	.96	.95	.87	.94	.84		
	76	.98	1.05	.96	.93	.96	.85	.95	.77	.96	1.11	.94	1.00	.95	1.14	.94	1.00		
	77	.98	1.05	.91	.49	.98	.97	.95	.99	.98	.73	.98	.84	.97	.85	.96	.79		
	78	.97	1.07	.96	1.00	.97	.57	.97	.72	.98	.78	.97	.70	.94	1.10	.93	1.00		
Mean			1.06		.81		.79		.85		.91		.90		.99		.94	0.94	0.88
High	75	.99	1.03	.98	.83	.96	.98	.95	1.00	.94	1.15	.92	1.00	.95	1.14	.94	1.00		
	76	.97	0.72	.96	.65	.96	1.09	.95	1.00	.98	.69	.97	.72	.93	1.19	.92	1.00		
	77	.97	1.06	.97	.99	.96	.63	.95	.72	.97	1.06	.96	.92	.95	1.13	.92	.91		
	78	.99	1.03	.97	.43	.94	.94	.94	.97	.96	1.08	.96	.94	.94	.95	.90	.78		
Mean			0.96		.74		.91		.94		1.00		.89		1.10		.94	0.99	0.88
Year means	75		1.05		0.85		0.85		0.90		1.01		0.99		1.01		0.96	0.98	0.92
	76		0.95		0.79		0.91		0.85		0.92		0.86		1.16		1.00	0.99	0.88
	77		1.06		0.80		0.80		0.89		0.95		0.91		1.03		0.89	0.96	0.86
	78		1.05		0.72		0.78		0.82		0.96		0.84		1.07		0.98	0.97	0.82
Pooled mean			1.03		0.80		0.83		0.87		0.96		0.89		1.07		0.97	0.97	0.88

$\hat{P}_i$ = maximum-likelihood estimate of the frequency of the ith allele in the outcrossed pollen pool

$\hat{t}$ = maximum-likelihood estimate of the outcrossing rate

generation result in outcrossing estimates greater than one. However, while not biologically meaningful, the conventional likelihood estimation (ML) procedure of Brown *et al.* (1975) will be less biased than the EM algorithm estimates when the true value of the outcrossing rate is approximately one. This arises because the EM algorithm, as presently formulated, is strictly bounded between zero and one in the estimate of the outcrossing rate. However, because the outcrossing rate (t) is calculated from classes of progeny sampled from a multinomial distribution, sampling variation can be expected to lead to estimates of t greater than one. Therefore, the restricted range of estimates obtained by the EM algorithm could lead to an underestimate of the effective outcrossing rate when the true value is near one in small samples. This bias will affect averaging of independent estimates to obtain average values. In an attempt to reduce this bias, the data from each of the sampling positions were pooled and the parameters re-estimated to obtain any of the pooled estimates for the crown strata, year or total summaries. Estimates were then averaged over loci.

A priori it is not expected that such potential for bias will destroy any differences among the various stratifications sampled in the population. Another discrepancy between the two procedures is in the data bases used to estimate the mating system parameters. If the maternal gamete contributed to each embryo is known, heterozygous embryos of the same genotype as the maternal

plant contribute information about the mating system for diallelic loci. Without the genetic information of the female gametophyte these embryos have to be omitted from the calculations. The Estimation-Maximization (EM) algorithm (Cheliak *et al.* 1982), which has been formulated specifically for gymnosperms, uses a slightly different data base to calculate mating system parameters than the method of Brown *et al.* (1975). With these qualifications and differences considered, the estimates of the mating system parameters from the EM algorithm will be used in all subsequent analyses.

The deviant comparisons (e.g., Middle 1977, and High 1978 for Aph-2) can be attributed to the heterozygous embryos from heterozygous maternal plants. For example, suppose there are two alleles i and j in the population and allele i is of considerably higher frequency than j. When a heterozygous maternal plant segregates for the i allele, an i allele will most likely be received upon outcrossing. Similarily, when the maternal plant segregates for the j allele, an i allele will most likely be received upon outcrossing. Thus, if the observed number of ij embryos is large when the maternal plant segregated for the i allele, the most likely event is that they have been derived from self-fertilization. The same argument applies to jj embryos when the maternal plant segregated for the j allele. These arguments necessarily assume a homogeneous distribution of allele frequencies in the pollen pool. Otherwise, estimates

of outcrossing will be biased downward. In this way, all of the data collected can contribute information to estimation of the mating system.

3.5.2 Mating system of jack pine

If the mating system is the only factor influencing the transmission of gametes from one generation to the next, all loci should give statistically comparable estimates for the apparent outcrossing rate in any particular sampling position. In general, there is considerable heterogeneity among the mating system parameters obtained from individual loci for any sampling position (Table 3.3).

3.5.3 Log-linear analyses

The heterogeneity of apparent outcrossing estimates (pooled over females) with respect to the distribution of the predicted numbers of selfed and outcrossed embryos was fitted by a hierarchy of log-linear logistic models. The simplest logistic model includes 'main effects' of locus, year and crown strata, but assumes all higher order interactions are nil. Even when all of the interaction terms among the independent variables were included, there was still significant heterogeneity among the estimated outcrossing rates (Table 3.4). However, 50% of the heterogeneity remaining after the fit of the baseline model could be accounted for by the penultimate logit model. Since the log-likelihood ratio statistic for goodness-of-fit can

Table 3.4. Anova of logit of expected numbers of outcrossed versus selfed progeny by log-linear models.

Source of Variation	df	G^2	Percent proportional decrease of total
Total	45	2228.7	-
Main Effects	8	486.1	22
Locus	3	345.7	16
Year	3	126.6	6
Stratum	2	13.8	1
All Pair-wise Effects	21	635.7	29
Locus x Stratum	6	232.6	10
Stratum x Year	6	212.9	10
Locus x Year	9	190.2	9
Remainder	13	1106.9	50

be decomposed additively, inferences can be made about the relative contribution of the various interactions to the reduction in total heterogeneity. Therefore, inter-locus heterogeneity appears to contribute the largest amount (16%) to the variability in outcrossing estimates, with years contributing about 6% and crown strata about 1% as main effects. All first-order interactions had nearly equal contributions of approximately 10%. Thus, these interactions make substantial contributions to the total heterogenity of outcrossing rates.

Ignoring the significant inter-locus heterogenity (G_3^2=556.9, where the subscript refers to the degrees of freedom), a maximum-likelihood G^2 analysis testing the hypothesis of homogeneity of the frequency of outcrossed and selfed embryos from each of the 12 sampling positions was rejected for each of the four loci (Table 3.5). Therefore, there is significant intra-locus heterogeneity among the 12 sampling positions pooled over the 30 females. To test the homogeneity of the stratifications made in the population, this same test was applied to the crown strata summaries pooled over years and to the year summaries pooled over the crown strata for each of the loci, ignoring the significant interactions of year by crown statum. In addition, an STP procedure was used to delineate homogeneous subsets within the data when significant heterogeneity existed (Sokal and Rohlf 1969).

Table 3.5. G^2 summary of homogeneity of outcrossing rates for single loci and pooled summaries.

Source	Locus			
	Aph-2	Mdh-3	Mdh-4	6Pgd-2
Total				
G^2	653.3	452.0	368.5	413.3
Years				
G^2	164.1	23.7	162.9	150.1
STP	5>6,7>8	5,7,>6>8	5>7>6,8	6>5,8>7
Stratum				
G^2	40.2	123.2	1.1	104.8
STP	L,M>H	H>M>L	H,M,L	L>M,H
	Pooled Summaries			
Loci				
G^2	556.9			
STP	6Pgd-2>Mdh-4,Mdh-3>Aph-2			
Years				
G^2	104.9			
STP	5>6,7>8			
Stratum				
G^2	.494			
STP	L,M,H			

> = Greater than : , = Not different than

L = Low, M = Middle and H = High

5 = 1975, 6 = 1976, 7 = 1977 and 8 = 1978

For the crown strata analyses there were no obvious general patterns of variation among the loci. Mdh-4 was the only locus where no differences were detected among the three levels within the reproductive canopy of the stand. Mdh-3 was the only locus which showed significant clinal patterns: the selfing rate increased significantly from 6% in the high stratum to 14% in the middle stratum and 21% in the lower crown stratum. At Aph-2, low and middle crown strata were not significantly different but had higher outcrossing rates than the high crown stratum. At 6Pgd-2, the low crown stratum again produced fewer selfed embryos than the middle and high crown strata, which were not different from each other.

Three loci (Aph-2, Mdh-3 and Mdh-4) indicated that significantly fewer apparently selfed embryos were produced in 1975 than in 1976, 1977 and 1978. As well, fewer apparently selfed embryos were produced in 1976 and 1977 than in 1978 for Aph-2 and Mdh-3. The effective selfing rates for these two years (1976 and 1977) were not significantly different. The effective outcrossing rate was significantly higher for 6Pgd-2 in 1976. At this locus, the effective outcrossing rate was not significantly different in 1975 and 1978 but there was a significantly lower selfing rate in both of these years than in 1977.

Assuming these loci are independent, the pooled distributions (over females and loci) for the three crown strata and four years were tested using the same G^2 and STP

procedures. While there is heterogeneity in the single-locus estimates, the pooled distributions indicated no significant differences among the three crown strata (Table 3.5). Thus, within any of these strata, an average of 12% of the zygotes measured at the viable embryo stage appear to have been derived from self-fertilization. In contrast, the pooled distributions for the years showed significant differences. Significantly fewer embryos appear to have been derived from self-fertilization in 1975 than in 1976, 1977 or 1978. The highest proportions of apparently selfed embryos were obtained from the most recent crop (1978). The effective selfing rates in 1976 and 1977 were not significantly different from each other. Thus, at this stage of the life cycle, about 9% of the zygotes from 1975 matings appear to have been derived from self-fertilization. The effective proportion of selfed embryos increased to about 12% in 1976, 14% in 1977 and 18% in 1978. Therefore, there appears to be a significant change in the mating system among the years sampled. Furthermore, there were no relationships among the outcrossing rates for the various years and heterozygosity of either the observed or outcrossed pollen pool.

3.5.4 Multilocus analysis

As discussed by Shaw *et al.* (1981) and Ritland and Jain (1981), multilocus estimators of mating system parameters are expected to be less sensitive than single locus estimators to effects which cause heterogeneity under mixed

mating systems. For example, consanguineous matings, as a result of limited pollen flow or family clustering, are expected to cause a downward bias in single locus estimators, whereas multilocus estimators are generally less sensitive to these types of non-random matings and to heterogeneity in the pollen pool. The multilocus estimate of the effective outcrossing rate (Shaw *et al.* 1981) for the filial population pooled over females, years and strata is 0.91. Thus, it appears that there is a possibility of restricted pollen flow coupled with some family structuring that gives rise to consanguineous matings. However, it does not appear that this is the major contribution to the apparent proportion of self-fertilization.

3.6 Pollen pool

3.6.1 Comparison of the mature and outcrossed allele frequencies

Using t-statistics to test for the equality of two percentages (Sokal and Rohlf 1969), few significant differences could be demonstrated in the frequency of the common allele between the mature population (i.e., maternal) and the outcrossed pollen pool (Table 3.6). The lack of difference can be attributed to the small sample size and large standard errors in the mature population. In general, there was less expected variability (i.e., h^2) in the

Table 3.6. t-test for the homogeneity of common allele frequencies observed in the mature population (female component) and the observed pollen pool (male component).

		Locus			
Stratum	Year	Aph-2	Mdh-3	Mdh-4	6Pgd-2
Low	1975	0.224	0.096	1.272	1.569
	1976	1.479	0.065	0.668	1.360
	1977	1.116	0.032	0.337	0.868
	1978	-	0.266	2.106*	1.223
Middle	1975	1.475	0.064	2.421*	1.418
	1976	1.189	0.032	0.938	1.367
	1977	2.582*	0.099	1.456	2.233*
	1978	0.861	0.096	1.311	1.163
High	1975	0.070	0.065	0.923	1.353
	1976	0.940	0.064	1.404	0.742
	1977	0.668	0.032	0.671	0.820
	1978	0.823	0.456	.0915	0.330
			Summaries		
Strata	Low	1.075	0.001	1.020	1.338
	Middle	1.585	0.214	1.187	1.559
	High	0.672	0.136	0.573	0.878
Years	1975	0.752	0.001	0.973	1.543
	1976	1.368	0.034	0.535	1.210
	1977	1.368	0.001	0.936	1.299
	1978	0.902	0.140	1.437	0.951
Total		1.140	0.035	0.920	1.272

* indicates significant at $\alpha = .05$

outcrossed pollen pool than in the mature population. The exception was Aph-2, where the rare allele had a higher frequency in the outcrossed pollen pool than in the mature population. Over all the loci, there was an average of approximately a 2% difference in the common allele frequency.

3.6.2 Homogeneity of the pollen pool

Two loci (Aph-2 and Mdh-4) showed significant inter-female heterogeneity for the number of heterozygous and homozygous embryos observed ($G^2_{1,}$=139.37 and $G^2_{1,}$=51.86, respectively). There were no significant differences among trees for this parameter at Mdh-3 and 6Pgd-2 ($G^2_{1,}$=24.59 and $G^2_{1,}$=26.54, respectively). These results suggest that: (1) there are heterogeneous distributions of pollen allele frequencies among offspring of females; (2) females incorporate non-identical alleles at different rates or (3) the probability of outcrossing (or conversely, self-fertilization) varies among the trees for Aph-2 and Mdh-4. The net effect of pollen pool heterogeneity is an expected upward bias of inbreeding coefficients and a downward bias in outcrossing estimates (Brown *et al.* 1975). From these results it is likely that the two low outcrossing estimates observed for Aph-2 are further confounded by heterogeneous pollen pool distributions.

3.7 Variance effective population numbers

The male variance effective population number was calculated from the following formula (Ewens 1979):

$$N_e = \frac{4(N_m N_f)}{N_m + N_f} \qquad -3-$$

where Ne is the effective population number, Nm is the number of males, and Nf is the number of females (Table 3.7). The number of females was assumed to be 30 for all calculations.

It has been further assumed that the variance effective number calculated by the methods of Yasuda (1969) is equivalent to the inbreeding effective number. However, this assumption is true only when the population size is nearly static, which implies that there are an average of two successful gametes per individual, and when the distribution of family sizes follows a Poisson distribution (Kimura and Crow 1963). The results presented in Table 3.7 are not intended to be absolute values. In fact, these results are dependent upon the absolute sample sizes in the mature and filial populations. They are intended simply as a measure of the extent of random drift in the particular sample being studied. For example, the extent of random drift in 1975 in the low crown stratum (based on the mature population) was the same as a population with a variance effective size of 42, with 30 females and 16 males, all with an equal chance to reproduce.

Table 3.7. Variance effective population numbers and variance effective male numbers calculated from the mature population and observed pollen pool.

1. Mature population allele frequencies.

	Population					Males				
	75	76	77	78	Hm	75	76	77	78	Hm
L	42	29	65	51	43	16	10	35	22	17
M	32	49	31	28	34	11	21	10	9	12
H	53	25	55	59	43	24	8	25	29	17
Hm	41	32	46	42	40	16	11	18	16	15

2. Observed pollen pool allele frequencies

	75	76	77	78	Hm	75	76	77	78	Hm
L	53	94	92	76	75	24	108	99	52	50
M	46	61	55	67	57	19	31	25	38	27
H	91	82	71	94	84	94	65	44	108	69
Hm	58	76	70	78	70	29	53	42	55	42

L,M,H refer to Low Middle and High Strata respectively.
75,76,77,78 refer to 1975 1976 1977 1978 respectively
Hm = Harmonic mean.

Based on the mature population, the harmonic mean indicated that the mean population sampled was equivalent to a population with a variance effective size of 40, with 30 females and 15 males, all equally participating in the mating for the filial generation. No obvious trends existed in either the year or crown strata summaries. However, the middle stratum and 1976 were quite different from all other summaries.

When we compare the results based on the mature population to those based on the observed pollen pool, the latter invariably gave larger variance effective numbers. As expected, the observed pollen pool takes into account migration as well as self-fertilization. Again, the middle stratum indicated a smaller variance effective number. However, there was considerably greater variation among the strata based on the observed pollen pool. In this case, 1975 was quite different from any of the other years.

Overall, this represents a 43% increase in the harmonic mean variance effective population number and a 64% increase in the variance effective number of males contributing to the filial generation. Based on the observed pollen pool the harmonic mean filial population is equivalent to a population with a variance effective number of 70, with 30 females and 42 males all randomly mating.

If these results are indicative of migration rates in natural populations of this species, they suggest that jack pine populations may exchange a considerable amount of

genetic information by way of migration. However, it may be misleading to attribute all of this increase strictly to migration. Since a larger variance effective number is the result of a smaller variance associated with two means, part of this increase could be due to self-fertilization or some level of consanguineous matings. For example, if only several males were involved in most of the fertilizations, there would be a smaller variance in the mean allele frequency in the filial population and the observed pollen pool than between the filial population and mature population. This results from a large proportion of full siblings in the filial generation, which would not affect the observed pollen pool allele frequencies. However, since mature population allele frequencies are calculated on the basis of observed genotypic distributions, they necessarily assume equal contribution and, therefore, would be expected to show a greater variance in means. Thus, if the dynamics of the mating system are such that not all individuals participate equally, the variance effective number will be inflated relative to the inbreeding effective number.

3.8 Discussion

The most likely factor responsible for excesses of homozygotes in the filial populations (as measured by Hardy-Weinberg equilibrium) is the mixed mating system. However, when the individual sampling positions which

indicated an excess of homozygotes (F>0) are compared to the expected proportion under inbreeding equilibrium (i.e., Fe=(1-t)/(1+t) (Fyfe and Bailey 1951) or Fe=s/(2-s) (Hayman 1953)), 53% indicated an excess of observed homozygotes and, 42% a deficiency after correction for the mating system (i.e., equilibrium inbreeding value). Situations where excesses are observed over predictions by the mating system can be the result of a variety of causes (Brown 1979). The most common explanation is the Wahlund effect (Wahlund 1928). This effect is the result of subdivision in the population with variance in allele frequency among the subdivisions. An excess of homozygotes results when the subdivided populations are considered as a single panmictic unit. Wahlund effects, however, are not the only factors which can result in excesses of homozygotes. Positive genetic assortative mating whereby individuals of like genotype tend to mate also will result in an excess of homozygotes (Crow and Felsenstein 1968). Similarly, this excess of homozygotes could be due to a restricted neighbourhood size with family structuring within the neighbourhoods (Levin and Kerster 1971, 1974). This model implies that while mating is random within the neighbourhood, mating individuals within the neighbourhood are on average more related genetically than random individuals from the population. Since the expected dispersal distance of seeds from these serotinous cones is limited, there is scope for family structures to develop in

natural populations. Given a genetic basis for phenological timing of flowering and no barriers to consanguineous matings, this model as well as the positive assortative mating model could satisfactorily account for the observed excesses of homozygotes after correction for the mating system. In addition, the parental generation from which the embryos have been sampled could be inbred. Thus, even if mating is strictly random in the parental generation, the filial generation will be more related than if it had been derived from unrelated, non-inbred parents (Squillace 1974). At this level of analysis then, all of these effects are confounded. With the present set of data, it is not clear if any or all of these models are operating, and to what degree each contributes to the results. The observed F distribution in the mature population suggests that the parental generation is not likely inbred. However, members of the parental generation could still be related. Apart from historical preference, the other factors cannot be separated.

Cases where a deficiency of homozygotes are observed can be due to a number of causes (Brown 1979). When the mating system is one of predominant outcrossing, such as that in jack pine, explanations requiring strong linkage disequilibrium are unlikely. Thus, heterozygous advantages for chromosomal segments including the marker loci and associative overdominance are commonly cited as explanations. The first model implies a general heterozygote

advantage (heterosis) either for the marker locus itself or for a block of genes with which it is in linkage disequilibrium. The second model, associative overdominance, is really a specialized case of the more general hitchhiking effect (Strobeck 1979). Strobeck (1979) has shown that if there is only a single locus at which there is a heterozygotic advantage, there is apparent selection not only at closely linked loci, but at any locus in the genome. As the number of advantageous heterozygotic loci increases, the apparent selection effect is expected to increase. Furthermore, this model applies to situations in which there is partial selfing, or a mixed mating system such as jack pine. Negative assortative mating also can lead to an excess of heterozygotes after correction for the mating system. This could be brought about in a number of ways. If the phenology of female receptivity on a tree is temporally different than the males on the same tree, and if males on genotypically different trees tend to coincide in anthesis with the receptivity of these females, negative assortative mating will result. That male and female flowers on the same tree mature at slightly different times has long been known as a general phenomenon in forest trees (Sarvas 1962). However, the validity of the second part of this argument is unknown. If there is competititon among male gametes mediated by maternal effects (for example through the pollination drop (Owens *et al.* 1981) such that gametes of dissimilar genotypes tended to be favored, heterozygote

excesses would result. This type of incompatibility mechanism is unknown in the Pinaceae. The primitive lethal-semi-lethal allele system that causes early embryo abortion in the homozygous state also will lead to heterozygote excess. Indirect observations of this system have been discussed in the previous chapter.

Differences in allele frequency between the male and female gamete pools also will result in excesses of heterozygotes after a correction for the mating system (Robertson 1965, Workman 1969). This potential cause for heterozygote excesses is generally ignored in monoecious plant populations, because each individual is expected to contribute equally to both the male and female gamete pools. As discussed earlier differential contribution and incorporation of gametes in the two pools is confounded with classical migration. Differences in allele frequency were evident in the male and female gamete pools in this study. However, because of large standard errors in the female pool, few of these differences were significant. Indirect evidence supporting both of these forms of migration can be obtained from detailed observations made in a seed orchard of *Pinus sylvestris* L. (Bhumibhamon 1978) and natural populations of *Picea abies* L. and *P. sylvestris* (Koski 1970). Results from the variance effective population size calculations also suggested that migration could be affecting the genetic composition of the filial populations.

In 1941, Fisher demonstrated that once a gene arises which permits self-fertility in an otherwise outcrossing species, this gene has an immediate evolutionary advantage. This advantage arises because more copies of this gene are passed on to the offspring by selfing than by outcrossing. Thus, over time, this gene is expected to increase in frequency and finally become fixed in the population. At this point, the species would be a predominant inbreeder. However, if this advantage were opposed by natural selection favoring outcrossed individuals (by some type of heterotic selection) the species would maintain its predominant outbreeding mating system. Previous work with jack pine has shown that a three- to four-fold reduction in the number of filled seed followed by approximately a three-fold reduction in germination (measured over a 30 day period) occurs after controlled self-fertilization (Rudolph 1976). This observation alone suggests that there may be some type of heterotic effect associated with outcrossed individuals relative to selfed individuals.

At this level of experimental control, all of the factors listed above could be responsible for the observed heterozygote excesses in the filial generation. If we are to isolate any of the factors, specific experiments will have to designed to separate these confounded effects. These experiments should involve controlled crosses as well as observations on flowering, phenology and effective transmission distances of male gametes.

In this study, estimates of mating system parameters for each of the four loci were made from a common set of embryos. Therefore, variation in the actual proportions of selfed and outcrossed embryos does not contribute to the large, single-locus heterogeneities observed. This variability must be due to either random variation (i.e., statistical noise) and/or violations in the assumptions implicitly made in the estimation procedure. Because of the compensation possible with many of the violations, only the pooled estimates will be discussed.

Contrary to initial expectations of differences among strata in apparent outcrossing rates and previous results (Fowler 1965), the crown strata did not show any differences in selfing rate. This could be due to the low density of the stand as well as the range in absolute height (5-22 m) and the active reproductive crown (5-18 m) among the individual trees. Variability in height implies that the low stratum for one tree could be equivalent to the high stratum for another. The low density of the trees in the stand would tend to keep the pollen pool well mixed, not allowing any stratification as might be expected in high density, uniform-height stands. While differences in selfing rates may exist among crown levels as a general feature in forest stands, they were not demonstrable with these data.

The significant temporal variation in the mating system can be accounted for by two related hypotheses. First, there could be real changes in the mating system among the years

and observations from this study are merely due to chance. However, these changes were independent of the heterozygosity of the pollen pool so that the linear temporal trend observed is not a trivial problem of spurious detectability of the EM estimator. The second related hypothesis of differential viability over time suggests that the probability of germination of a selfed or inbreed zygote decreases over time. This same trend was reported by Moran and Brown (1980). They calculated the probability that a selfed embryo of alpine ash viable at time t would be viable at time t+2 years to be 0.56, if survivorship was due only to differential viability. Using this same strategy, these data suggest that the probability that a viable selfed embryo produced in 1975 is still viable in 1978 is 0.56.

Apparently dormant seed can deteriorate over time, particularly under sub-optimal conditions for long-term storage (Abdul-Baki and Anderson 1972). These data suggest that a greater selective reduction of selfed relative to outcrossed embryos, coupled with changes in the mating system, could occur over time. However, studies on mating system dynamics and 'dormant' embryo physiology are necessary to gain a better understanding of temporal variation in the mating system of serotinous-coned species. In addition, independent samples of the crop, assessing different ranges in years, should be made to determine if the same trends are demonstrable.

The observations of Moran and Brown (1980) and this study suggest that to maximize the number of unique genomes for domestication programs or gene conservation, sampling strategies should involve collection preferentially from older seed-crops. However, as there is a general decline in viability of jack pine seeds over time (Baker 1980) as well as an increased probability of insect attack, definition of the optimal strategy needs to account for other biotic factors.

Using morphological markers, Fowler (1965), Rudolph (1979) and Sittman and Taylor (1971) estimated selfing rates in jack pine populations to be 19%, 7% and 10%, respectively. These estimates compare favorably with those derived from enzyme markers in this study. In fact, the means are essentially the same (88% outcrossing). If this stand were to be burned and left to regenerate naturally, we would expect an average of 12% of the zygotes to be the direct result of apparent self-fertilization, if only the 1975 to 1978 crops were included. However, because of severe inbreeding depression associated with selfs in jack pine (Rudolph 1981), most of the selfed seedlings are not expected to survive to the next mature population. This contention is further supported by heterozygote excesses observed not only in jack pine, but in most mature conifers (Linhart *et al.* 1981, O'Malley *et al.* 1979, Yeh and Layton 1979). If there is selective removal of selfed individuals, it would be interesting to consider the potential

contributions to the subsequent mature population from each of the years of mating in terms of the life cycle and mating system components (Allard *et al.* 1977 and Clegg *et al.* 1978).

3.9 Literature Cited

Abdul-Baki, A.A., and J.D. Anderson. 1972. Physiological and biochemical deterioration of seeds. p 283-315 in Seed biology. vol 2. (T.T. Kozlowski ed.) Academic Press, N.Y.

Allard, R.W., A.L. Kahler and M.T. Clegg. 1977. Estimation of mating cycle components of selection in plants. p 1-19 in Measuring selection in natural populations. (F.B. Christiansen and T.M. Fechenel, eds.) Springer-Verlag Berlin.

Baker, W.D. 1980. Jack pine cone and seed quantity and quality from Gogama, Ontario. M.Sc. thesis, Univ. of Toronto.

Bhumibhamon, S. 1978. Studies on Scots pine seed orchards in Finland with special emphasis on the genetic composition of the seed. Comm. Inst. For. Fenn. 94(4):1-118.

Brown, A.H.D. 1979. Enzyme polymorphisms in plant populations. Theor. Pop. Biol. 15:1-42.

Brown, A.H.D., A.C. Matheson and K.G. Eldridge. 1975. Estimation of the mating system of *Euclayptus obliqua* L'Herit. using allozyme polymorphisms. Aust. J. Bot.

23:931-949.

Cheliak, W.M., K. Morgan, C. Strobeck, F.C.H. Yeh and B.P. Dancik. 1982. Estimation of mating system parameters in plant populations using the EM algorithm. (Submitted Theor. Appl. Genet.)

Clegg, M.T., A.L. Kahler and R.W. Allard. 1978. Estimation of life cycle components of selection in an experimental plant population. Genetics 89:765-792.

Crow, J.F. 1954. Breeding structure of populations. II. Effective population number. p 543-556 in Statistics and mathematics in biology (O. Kempthorne, T.A. Bancroft, J.W. Gowen and J.L. Lush eds.). Iowa State College Press, Iowa.

Crow, J.F., and J. Felsenstein. 1968. The effect of assortative mating on the genetic composition of a population. Eugen. Quart. 15:85-97.

Crow, J.F., and M. Kimura. 1970. An introduction to population genetics theory. Burgess Press, Minneapolis, Minnesota.

Ewens, W.J. 1979. Mathematical population genetics. Springer-Verlag, N.Y.

Fisher, R.A. 1941. Average excess and average effect of a gene substitution. Ann. Eugen. 11:53-63.

Fowler, D.P. 1965. Natural self-fertilization in three jack pines and its implications in seed orchard management. For. Sci. 11:56-58.

Fyfe, J.L., and N.T.J. Bailey. 1951. Plant breeding studies

in *Leguminous* forage crops. I. Natural cross-breeding in winter beans. J. Agri. Sci. 41:371-378.

Hamrick, J.L., Y.B. Linhart and J.B. Mitton. 1979. Relationships between life history characteristics and electrophoretically-detectable genetic variation in plants Ann. Rev. Ecol. Syst. 10:173-200.

Hayman, B.I. 1953. Mixed selfing and random mating when heterozygotes are at a disadvantage. Heredity 7:185-192.

Kimura, M., and J.F. Crow. 1963. The measurement of effective population number. Evol. 17:279-288.

Koski, V. 1970. A study of pollen dispersal as a mechanism of gene flow in conifers. Comm. Inst. For. Fenn. 70(4):1-78.

Levene, H. 1949. On a matching problem arising in genetics. Ann. Math. Statist. 20:91-94.

Levin, D.A., and H.W. Kerster. 1971. Neighbourhood structure in plants under diverse reproductive methods. Amer. Nat. 105:345-354.

Levin, D.A., and H.W. Kerster. 1974. Gene flow in seed plants. Evol. Biol. 7:139-220.

Linhart, Y.B., J.B. Mitton, K.B. Sturgeon and M.L. Davis. 1981. Genetic variation in space and time in a population of ponderosa pine. Heredity 46:407-426.

Malecot, G. 1969. The mathematics of heredity. W.H. Freeman, San Francisco.

Moran, G.F., and A.H.D. Brown. 1980. Temporal heterogeneity of outcrossing rates in alpine ash (*Eucalyptus*

delegatensis R.T. Bak.) Theor. Appl. Genet. 57:101-105.

O'Malley, D.M., F.W. Allendorf and G.M. Blake. 1979. Inheritance of isozyme variation and heterogeneity in *Pinus ponderosa*. Biochem. Genet. 17:233-250.

Owens, J.N., S.J. Simpson and M. Molder. 1981. Sexual reproduction of *Pinus contorta*. I. Pollen development, the pollination mechanism and early ovule development. Can. J. Bot. 59:1828-1843.

Ritland, K., and S.K. Jain. 1981. A model for the estimation of outcrossing rate and gene frequencies using *n* independent loci. Heredity 47:35-52.

Robertson, A. 1965. The interpretation of genotypic ratios in domestic animal populations. Anim. Prod. 7:319-324.

Rudolph, T.D. 1976. Cone set, seed yield, seed quality and early seeding development of S_2 generation jack pine. p 42-60 in Proc. Tenth Cent. States For. Tree Imp. Conf.

Rudolph, T.D. 1979. Seed production in the first eight years and frequency of natural selfing in a simulated jack pine seed orchard. p 33-47 in Proc Thirteenth Lake States For. Tree Imp. Conf.

Rudolph, T.D. 1981. Four year height growth variation among and within S_0, $S_1 x S_1$, S_1 open-pollinated and S_2 jack pine families. Can. J. For. Res. 11:654-661.

Sarvas, R. 1962. Investigations on the flowering and seed crop of *Pinus silvestris*. Comm. Inst. For. Fenn. 53(4):1-198.

Shaw, D.V., A.L. Kahler and R.W. Allard. 1981. A multilocus

estimator of mating system parameters in plant populations. Proc. Nat. Acad. Sci. (USA) 78:1298-1302.

Sittman, K., and H. Tyson. 1971. Estimates of inbreeding in *Pinus banksiana*. Can. J. Bot. 49:1241-1245.

Sokal, R.R., and F.J. Rohlf. 1969. Biometry. W.H. Freeman, San Francisco.

Squillace, A.E. 1974. Average genetic correlations among offspring from open-pollinated forest trees. Silvae Genetica 23:149-156.

Strobeck, C. 1979. Partial selfing and linkage: The effect of a heterotic locus on a neutral locus. Genetics 92:305-315.

Wahlund, S. 1928. Zammensetzung von populationen und Korrelationsershienungen von Standpunkt aus befrachtet. Hereditas 11:65-106.

Workman, P.L. 1969. The analysis of simple genetic polymorphisms. Hum. Biol. 41:97-114.

Wright, S. 1965. The interpretation of population structure with special regard to systems of mating. Evol. 19:395-420.

Yasuda, N. 1969. The estimation of the variance effective population number based on gene frequency. Jap. Jour. Human Genet. 14:10-16.

Yeh, F.C., and C. Layton. 1979. The organization of genetic variability in central and marginal populations lodgepole pine, *Pinus contorta* ssp. *latifolia*. Can. J. Genet. Cytol. 21:487-503.

4. PRACTICAL IMPLICATIONS

4.1 Introduction

In most studies dealing with forest trees, it is assumed that open-pollinated progeny are the result of random mating with an effectively infinite male gamete pool. This is equivalent to assuming that open-pollinated progeny are half-siblings from random mating (Wright 1976). These assumptions allow simplifications to be made when drawing inferences from the performance of open-pollinated progeny and in estimation procedures for genetic parameters such as additive genetic variance. However, results from this study have demonstrated that the mixed mating system model as well as a finite effective male gamete pool are more realistic assumptions. Even after correction for the mating system it is clear that mating was not completely random for any of the four loci studied. There are several practical implications of these results.

4.1.1 Inference

If we accept the results from this study -- finite effective male populations, temporally and spatially variable partial-selfing rates and evidence for non-random mating -- the performance of open-pollinated progeny can be interpreted in two ways. Typically, genetic differences are inferred when a provenance or family mean is significantly

different from the test mean. The inferred genetic differences instead may be a reflection of the mating system. Severe inbreeding depression associated with consanguineous matings in conifers (Hadders and Koski 1975) suggests that the inferred genetic differences could be confounded by the specific mating event represented in a particular open-pollinated provenance or family sample. Thus, differences in performance could be in part the result of different selfing rates, different male effective sizes or different patterns of non-random mating among the provenances or open-pollinated families sampled. With knowledge of the mating system it is possible to identify, at least at the population or provenance level, differences in mating systems in the particular samples representing those populations. Specific field performance or ranking of family or provenance means can be interpreted in light of expected inbreeding depression if estimates of the mating system are available.

4.1.2 Estimation

When estimates of additive genetic variance are made from open-pollinated progeny it is usually assumed that these individuals are half-siblings. Strictly speaking, however, half-siblings are individuals that have one parent in common and the other parent different (Falconer 1961). A group or family of half-siblings, therefore, represents the progeny of one maternal plant mated at random and producing

only one offspring by each unrelated pollen source. Therefore, for half-siblings to obtain, there must be an effectively infinite pollen pool relative to the number of seeds collected from a single maternal plant. Two biases can affect the estimates of genetic parameters if these initial assumptions are incorrect. The first bias, a mixed mating system, can be corrected as shown in Figure 4.1. This figure has been constructed from the following formula:

$$K^{-1} = \left(\frac{1}{2}\left(\frac{(1 + F)}{2}\right)\right)^{-1} \qquad -1-$$

where F is the correlation between uniting gametes. For true half-siblings the correlation between uniting gametes is 0 and the expected coefficient of additive genetic variance is 4.0. However, when the level of family structure (i.e., inbreeding relative to assuming half sibs) increases to complete full siblings, the value of F becomes 1 and the expected coefficient of additive genetic variance is 2.0 Thus, this correction measures the effect of inbreeding on the expected coefficient of additive genetic variance for half-sibling families. The value of F in Figure 4.1 is assumed to be the equilibrium inbreeding value (i.e., Fe=(1-t)/(1+t) (Fyfe and Bailey 1951) or Fe=s/(2-s) (Hayman 1953)). With the outcrossing results presented in this study, the expected coefficient of additive genetic variance varies from 3.82 for 1975 to 3.67 for 1978 accounting for inbreeding alone (Table 4.1). These biases from inbreeding

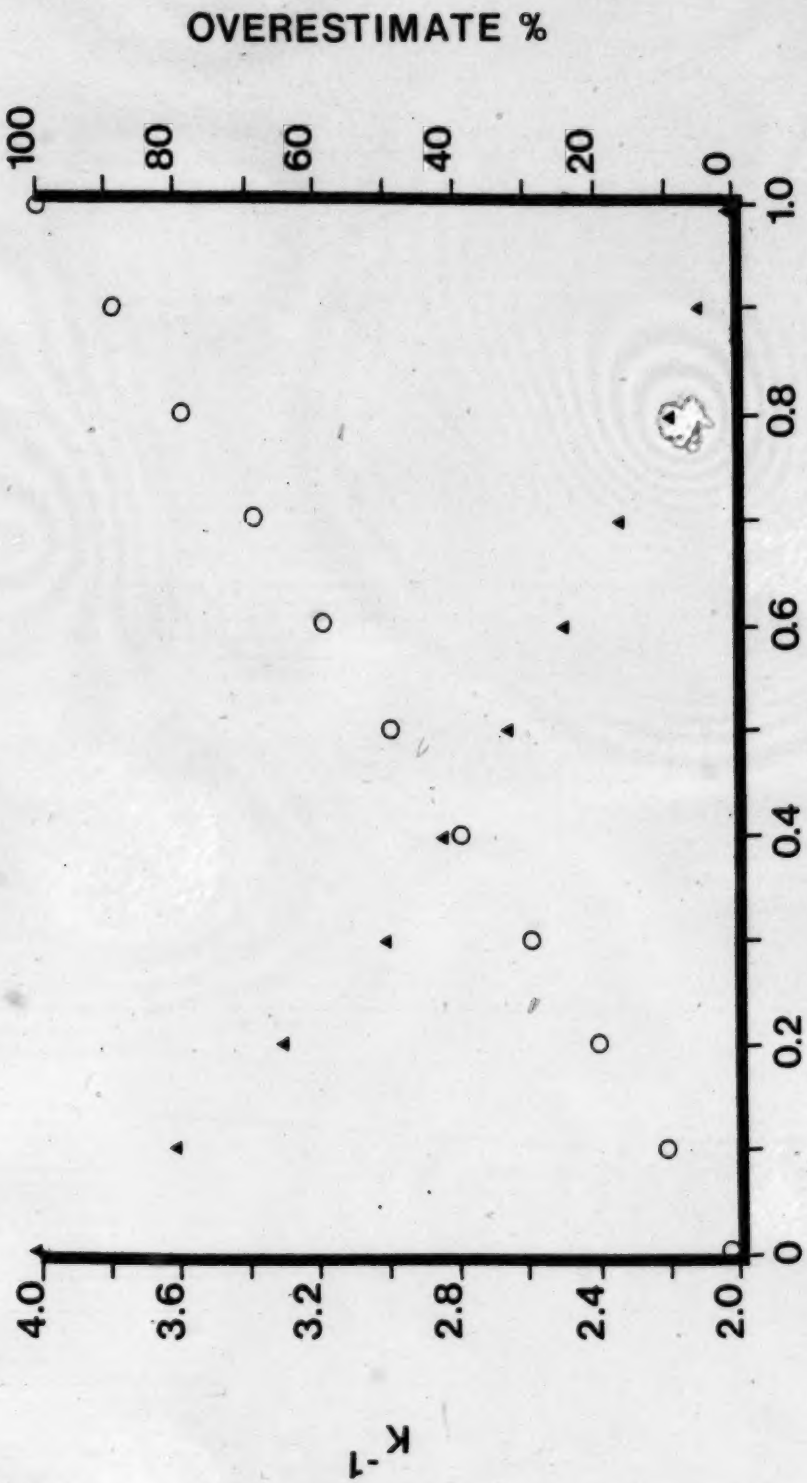

Figure 4.1 Expected coefficient of additive genetic variance (K^{-1}) with various levels in inbreeding (F).

Table 4.1. Expected coefficients of additive genetic variance under inbreeding and various family structures for the estimated mating system of a stand of jack pine.

			K^{-1}		Overestimate	
Year	t	Fe	hs	fs	hs	fs
1975	0.92	0.047	3.82	2.88	4.5	28.0
1976	0.88	0.062	3.77	2.84	5.8	29.0
1977	0.86	0.074	3.72	2.80	7.0	30.0
1978	0.82	0.094	3.67	2.75	8.9	31.3
Mean	0.88	0.067	3.75	2.82	6.3	29.5

t=outcrossing rate, Fe=equilibrium inbreeding value
hs=half-sibling family, fs=half+full-sibling families
K^{-1}=expected coefficient of additive genetic variance

represent an average overestimate of the total additive genetic variance of 6 percent. However, it is still assumed that the proportion of progeny derived from outcrossing represents half-sibling families. If the progeny sampled actually represent collections of half- and full-sibling families, which is expected with a finite male pool, a formula presented by Yeh (1982) can be used to estimate the expected coefficent of additive genetic variance for varying degrees of inbreeding structure and number of seedlings per parent tree. Clearly, the degree of bias due to the assumption that all families represent half-siblings is dependent on not only the effective male population size, but also the sample size from a particular female and F (the inbreeding coefficient). If we accept 42 as the estimate of the male effective size and that 120 zygotes are sampled from each of 30 females, we would expect a maximum of 41 full-sibling families each with a group size of approximately three. With the levels of inbreeding observed and this degree of family structure in the filial population, the expected coefficient of additive genetic variance ranges from 2.88 for 1975 to 2.75 for 1978 (Table 4.1). If we had assumed the sample represented half-sibling families with no inbreeding, the total additive genetic variance would have been overestimated by 29.5 percent.

Thus, even though the level of inbreeding appears to be small and the effective male population appears to be large, the assumptions of random mating and half-sibling families

can result in considerable bias when estimating genetic parameters. Furthermore, these adjustments do not account for all the bias. Within each of the full-sibling families, no account has been made of non-additive effects of dominance, epistasis and higher-order interactions. However, without knowledge of the pedigree it is impossible for us to adjust for these additional sources of bias.

4.2 Seed Orchards

Since there should be no family structure among the trees in a seed orchard, it is expected that the seeds produced will be free of any inbreeding apart from selfing. There are two sources of selfing in seed orchards. The first type is the same as that expected in natural populations, i.e., self-matings within the tree. However, selfed seeds also can be derived from "outcrossed" fertilizations among ramets of the same clone within the seed orchard. Since no data are available, it is difficult to estimate the relative proportions of these two types of selfed seeds produced in seed orchards. It is likely that the proportions of this latter type of selfing will be functionally related to the absolute distance among ramets of the same clones, as well as a number of environmental parameters (Shen *et al.* 1981). Furthermore, it is usually assumed that the seeds derived from seed orchards are the result of panmixia. With the results obtained from this study, this assumption is likely

invalid. If pollen dispersion is limited both spatially as well as temporally, as can be expected with collections of diverse material flowering at different times, some type of management may be necessary to ensure that the effective sizes of the population remain as large as possible. These management techniques may involve booster pollination, design of seed orchards to maximize mixing of the pollen pool, and overhead irrigation to delay and synchronize flowering. However, quantitative estimates of the mating system and its dynamics in the seed orchard should be made before embarking on an expensive management program. If some type of management program is implemented its effectiveness should be constantly monitored by re-evaluating estimates of the mating system and its dynamics.

4.2.1 Current nursery,greenhouse and plantation practices

In nursery practice and particularly in container production of seedlings, attempts are made to provide optimal environments for germination and growth. Competition among plants is carefully controlled by optimal density/spacing regimes in nursery beds. In containerized production systems, competition effectively has been eliminated. In intensive forest management, optimal spacing/density regimes, nutrition, and protection against losses from competition, insect and disease attacks are provided through artificial regeneration and cultural management programs. Minimization of regeneration costs is

the incentive for these programs. With reduced levels of competition it is likely that the proportion of selfed individuals maintained in the population will increase. Under natural conditions, these individuals would have been removed by competition and/or natural selection; under these controlled conditions, they will not. Therefore, it is a critical responsibility for forest geneticists to ensure that the maximum proportion of seeds derived from seed orchards or seed production areas are the result of outcrossing. From quantitative studies of mating systems and the dynamics of mating systems in natural populations, seed production areas and seed orchards, strategies can be formulated to ensure that expected gains from tree improvement are not compromised.

4.3 Literature Cited

Falconer, D.S. 1961. Introduction to quantitative genetics. Ronald Press, N.Y.

Fyfe, J.L., and N.T.J. Bailey. 1951. Plant breeding studies in *Leguminous* forage crops. I. Natural cross-breeding in winter beans. J. Agri. Sci. 41:371-378.

Hadders, G., and V. Koski. 1975. Probability of inbreeding in seed orchards. p 108-117 in Seed orchards (R. Faulkner ed.) For. Comm. Bull. 54. H.M.S.O., London.

Hayman, B.I. 1953. Mixed selfing and random mating when heterozygotes are at a disadvantage. Heredity 7:185-192.

Wright, J.W. 1976. Introduction to forest genetics. Academic Press, N.Y.

Shen, H-H, D. Rudin and D. Lindgren. 1981. Study of the pollination pattern in a Scots pine seed orchard by means of isozyme analysis. Silvae Genetica 30: 7 - 15.

Yeh, F.C.H. 1982. Expected bias in open-pollinated families with inbreeding, selfing and family structure (in preparation).

5. APPENDIX 1.

Variance effective population estimation method

The method used to estimate the variance effective population number in the filial population from Yasuda (1969) is as follows.

Let the frequency of an allele (A_i) at a locus in the mature population be $p_{i(m)}$ and the corresponding frequency of that allele in the filial population be $p_{i(f)}$, where $p_{i(f)} = o_i/2N$, o_i is the number of A_i alleles observed in the filial population and 2N is the number of gametes which have been contributed to the filial generation. These 2N gametes are considered as a random sample from an infinite pool of gametes produced by the parents. Since o_i is a random variable which is sampled from a binomial distribution the mean [$p_{i(f)}$] and variance [$V_{p(f)}$] of o_i are given by:

$$p_{i(f)} = p_{i(m)} \qquad \text{and} \qquad V_{p(f)} = p_{i(m)}(1-p_{i(m)})/2N$$

The change in allele frequency between the two generations, $\delta p = p_{i(f)} - p_{i(m)}$ is expected to be 0, but this quantity has a variance of $V_{\delta p} = p_{i(m)}(1-p_{i(m)})/2N$. To normalize the variance and make it in dependent of the allele frequencies, the data are transformed by

$$p = \sin^2\Theta$$

or $$\Theta = \sin^{-1}\sqrt{p}.$$

Letting $\delta\Theta = \Theta_m - \Theta_f$, the variance of $\delta\Theta$ is

$$V_{\delta\Theta} = \left(\left[\frac{\partial\Theta}{\partial p_{(m)}}\right]^2 V_{\delta p_m}\right) = \frac{1}{8N}$$

which can be computed by

$$nV_{\delta\Theta} = \frac{1}{2} \sum_{k_2} (\delta\Theta_1^2 + \delta\Theta_1^2) + \frac{2}{3} \sum_{k_3} (\delta\Theta_1^2 + \delta\Theta_2^2 + \delta\Theta_3^2) + \ldots.$$

$$.. + \frac{(m-1)}{m} \sum_{k_m} (\delta\Theta_1^2 + \ldots. + \delta\Theta_m^2),$$

where $n = k_2 + 2k_3 + \ldots.. + (m-1)k_m$ is the number of independent alleles observed and k_i is the number of loci with i alleles. The variance effective number is then given by

$$N_e = \frac{1}{8V_{\delta\Theta}} .$$

Literature Cited

Yasuda, N. 1969. The estimation of variance effective population number based on gene frequency. Jap. J. Hum. Genet. 14: 10-16.

6. APPENDIX 2.

EM algorithm to estimate mating systems in gymnosperms

To use the EM algorithm to estimate mating system parameters in gymnosperms the following data are necessary:

1. A known or inferred maternal genotype. These data can be obtained from either a pedigree or by inference from analysis of the haploid megagametophytic tissue.
2. Genotypes of open-pollinated progeny arrays as well as known maternal contribution to heterozygous embryos from known or inferred heterozygous maternal plants of the same genotype.
3. At least two different classes of maternal genotypes

The notation for this formulation is consistent with that adopted in Cheliak *et al.* 1982.

6.1 The expectation step (E)

Table A2.1 outlines the phenotypic classes observed after various mating events, conditional on the maternal genotype. Given that we have observed a total of X embryos from Y maternal plants, the expected number of selfed (..x..) and outcrossed (..x:) embryos in this genotypic class are given by equations 1a and 1b, for homozygous and heterozygous maternal plants, respectively.

In the case of heterozygous maternal plants a distinction can be made between embryos which received an i or j maternal allele when the paternal gamete is the same as one of the two alleles of the maternal plant.

Table A2.1 Observed phenotypes and underlying genotypes resulting from various mating events.

Maternal Genotype	Progeny Genotype	Incomplete Space (observed numbers)	Probability	Complete Space (expected numbers)	
$A_i^*A_i^*$	$A_i^*A_i^*$	$ii^y{}_{ii}$	s	$ii^x{}_{ii}$	
	$A_i^*A_i$		$(1-s)p_i$	$ii^i_x{}_i$	
	$A_i^*A_j$	$ii^y{}_{ij}$	$(1-s)p_j$	$ii^j_x{}_i$	$j\neq i$
$A_i^*A_j^*$	$A_i^*A_i^*$	$ij^y{}_{ii}$	$s/4$	$ij^x{}_{ii}$	
$j\neq i$	$A_i^*A_i$		$(1-s)p_i/2$	$ij^i_x{}_i$	
	$A_i^*A_j^*$	$ij^y{}_{ij}$	$s/4$	$ij^x{}_{ij}$	
	$A_i^*A_j$		$(1-s)p_j/2$	$ij^j_x{}_i$	
	$A_j^*A_i^*$	$ij^y{}_{ji}$	$s/4$	$ij^x{}_{ji}$	
	$A_j^*A_i$		$(1-s)p_i/2$	$ij^i_x{}_j$	
	$A_j^*A_j^*$	$ij^y{}_{jj}$	$s/4$	$ij^x{}_{jj}$	
	$A_j^*A_j$		$(1-s)p_j/2$	$ij^j_x{}_j$	
	$A_i^*A_\ell$	$ij^y{}_{i\ell}$	$(1-s)p_\ell/2$	$ij^\ell_x{}_i$	$\ell\neq i,j$
	$A_j^*A_\ell$	$ij^y{}_{j\ell}$	$(1-s)p_\ell/2$	$ij^\ell_x{}_j$	$\ell\neq i,j$

$ii^{x}_{ii} = ii^{y}_{ii}$	$s/[s + p_i(1-s)]$		1-a
$ii^{x}_{i}{}^{i} = ii^{y}_{ii}$	$p_i(1-s)/[s + p_i(1-s)]$		
$ii^{x}_{i}{}^{j} = ii^{y}_{ij}$		$j \neq i$	
$ij^{x}_{ii} = ij^{y}_{ii}$	$s/[s + 2p_i(1-s)]$	$j \neq i$	1-b
$ij^{x}_{i}{}^{i} = ii^{y}_{ii}$	$2p_i(1-s)/[s + 2p_i(1-s)]$		
$ij^{x}_{ij} = ij^{y}_{ij}$	$s/[s + 2p_j(1-s)]$	$j \neq i$	
$ij^{x}_{i}{}^{j} = ij^{y}_{ij}$	$2p_j(1-s)/[s + 2p_j(1-s)]$		
$ij^{x}_{ji} = ij^{y}_{ji}$	$s/[s + 2p_i(1-s)]$	$j \neq i$	
$ij^{x}_{j}{}^{i} = ij^{y}_{ji}$	$2p_i(1-s)/[s + 2p_i(1-s)]$		
$ij^{x}_{jj} = ij^{y}_{jj}$	$s/[s + 2p_j(1-s)]$	$j \neq i$	
$ij^{x}_{j}{}^{j} = ij^{y}_{jj}$	$2p_i(1-s)/[s + 2p_i(1-s)]$		
$ij^{x}_{i}{}^{\ell} = ij^{y}_{i\ell}$		$\ell \neq j \neq i$	
$ij^{x}_{j}{}^{\ell} = ij^{y}_{j\ell}$			

Throughout 1-a and 1-b, p_i, p_j, p_ℓ are the frequencies of the ith, jth and ℓth alleles in the outcrossed pollen pool, i, j, ℓ = 1, ..., r, and s is the selfing rate.

6.1.1 The maximization step (M)

Maximum-likelihood estimators of p and s now can be derived simultaneously from the complete sample space (X) using conventional maximum-likelihood theory (Dempster *et al.* 1977). The likelihood equation for the complete set of observations is given by:

$$L = N \prod_{i=1}^{k} (s)^{ii x_{ii}} (p_i(1-s))^{ii x_i^i} \prod_{j \neq i}^{k} (p_j(1-s))^{ii x_i^j}$$

$$\prod_{i=1}^{k} \prod_{j=i+1}^{k} (s/4)^{ij x_{ii}} (p_i(1-s)/2)^{ij x_i^i} (s/4)^{ij x_{ij}} (p_j(1-s)/2)^{ij x_i^j}$$

$$(s/4)^{ij x_{ji}} (p_i(1-s))^{ij x_j^i} (s/4)^{ij x_{jj}} (p_i(1-s))^{ij x_j^j}$$

$$\prod_{l \neq i,j}^{k} (p_l(1-s)/2)^{ij x_i^l} (p_l(1-s)/2)^{ij x_j^l}$$

where N is a normalizing constant, p(i), i=1,..,k is the frequency of the i th allele in the outcrossed pollen pool, s is the selfing rate and x is an expected observation. Let the log likelihood equation be:

$$L^* = \ln (L)$$

Using the method of Lagrangian multipliers to constrain the allele frequencies to sum to 1.0, the likelihood equation is:

$$L^{\dagger} = L^* + \lambda \left(1 - \sum_{i=1}^{k} p_i\right) ,$$

where the Lagrangian constraint (λ) is:

$$\sum_{i=1}^{k} p_i = 1 \ .$$

Taking partial differentials of L† with respect to p(i), i=1,..,k, s and λ, equating them to 0 and solving for λ, we obtain the following maximum-likelihood estimators for p_r and s:

$$\hat{p}_r = \frac{x^r_{\cdot\cdot\cdot}}{x^{\cdot}_{\cdot\cdot\cdot}}$$

and

$$\hat{s} = \frac{x_{\cdot\cdot\cdot\cdot}}{x_{\cdot\cdot\cdot\cdot} + x^{r}_{\cdot\cdot\cdot}}$$

where,

$$x^r_{\cdot\cdot\cdot} = \sum_{i=1}^{k} {}_{ii}x^r_i + \sum_{j=i+1}^{k} {}_{ij}x^r_i + {}_{ij}x^r_i$$

$$x^{\cdot}_{\cdot\cdot\cdot} = \sum_{r=1}^{k} x^r_{\cdot\cdot\cdot}$$

$$x_{\cdot\cdot\cdot\cdot} = \sum_{i=1}^{k} {}_{ii}x_{ii} + \sum_{j=i+1}^{k} {}_{ij}x_{ii} + {}_{ij}x_{ij} + {}_{ij}x_{ji} + {}_{ij}x_{jj}$$

6.2 Literature cited

Cheliak, W.M., K. Morgan, C. Strobeck, F.C.H. Yeh and B.P. Dancik. 1982. Estimation of mating system parameters in plant populations using the EM algorithm. (Submitted Theor. Appl. Genet.)

Dempster, A.P., N.M. Laird, D.B. Rubin. 1977. Maximum likelihood from incomplete data via the EM algorithm. J. Royal Stat. Soc. B. 39:1-38.